JMIR Human Factors

Impact Factor (2022): 2.7
Volume 11 (2024) ISSN 2292-9495 Editor in Chief: Andre Kushniruk, BA, MSc, PhD, FACMI

Contents

Original Papers

Acceptance and User Experiences of a Wearable Device for the Management of Hospitalized Patients in COVID-19–Designated Wards in Ho Chi Minh City, Vietnam: Action Learning Project (e44619)
An Luu, Truong Nguyen, Van Cao, Trinh Ha, Lien Chung, Trung Truong, Tung Nguyen Le Nhu, Khoa Dao, Hao Nguyen, Phan Khanh, Khanh Le, Luu Tran, Phung Nhat, Duc Tran, Yen Lam, Catherine Thwaites, Jacob Mcknight, Nguyen Vinh Chau, Jennifer Van Nuil, Vietnam ICU Translational Applications Laboratory (VITAL). 9

Exploring a Gaming-Based Intervention for Unemployed Young Adults: Thematic Analysis (e44423)
Ingjerd Straand, Asbjørn Følstad, Jone Bjørnestad. 19

Evaluating the User Experience of a Smartphone-Delivered Sexual Health Promotion Program for Older Adults in the Netherlands: Single-Arm Pilot Study (e56206)
Ana Correia de Barros, Mariëtte Bergmans, Kreshnik Hasanaj, Drianë Krasniqi, Catarina Nóbrega, Bruna Carvalho Carneiro, Priscila Vasconcelos, Ana Quinta-Gomes, Pedro Nobre, Joana Couto da Silva, Cristina Mendes-Santos. 39

Designing for Improved Patient Experiences in Home Dialysis: Usability and User Experience Findings From User-Based Evaluation Study With Patients With Chronic Conditions (e53691)
Anna Aspelund, Paula Valkonen, Johanna Viitanen, Virpi Rauta. 54

Enabling Health Information Recommendation Using Crowdsourced Refinement in Web-Based Health Information Applications: User-Centered Design Approach and EndoZone Informatics Case Study (e52027)
Wenhao Li, Rebecca O'Hara, M Hull, Helen Slater, Diksha Sirohi, Melissa Parker, Niranjan Bidargaddi. 65

Participatory Development of an Integrated, eHealth-Supported, Educational Care Pathway (Diabetes Box) for People With Type 2 Diabetes: Development and Usability Study (e45055)
Daan de Frel, Mariëlle Schroijen, Jiska Aardoom, Wesley van Gils, Sasja Huisman, Veronica Janssen, Anke Versluis, Maaike Kleinsmann, Douwe Atsma, Hanno Pijl. 80

Assessing the Utility, Impact, and Adoption Challenges of an Artificial Intelligence–Enabled Prescription Advisory Tool for Type 2 Diabetes Management: Qualitative Study (e50939)
Sungwon Yoon, Hendra Goh, Phong Lee, Hong Tan, Ming Teh, Dawn Lim, Ann Kwee, Chandran Suresh, David Carmody, Du Swee, Sarah Tan, Andy Wong, Charlotte Choo, Zongwen Wee, Yong Bee. 96

Preferences, Needs, and Values of Patients With Chronic Obstructive Pulmonary Disease Attending a Telehealth Service: Qualitative Interview Study (e53131)
Camilla Schmidt, Karen Borgnakke, Anne Frølich, Lars Kayser. 107

Predictive Factors of Physicians' Satisfaction and Quality of Work Under Teleconsultation Conditions: Structural Equation Analysis (e47810)
Liliana Hawrysz, Magdalena Kludacz-Alessandri, Renata Walczak. 121

XSL•FO
RenderX

Safety in Teletriage by Nurses and Physicians in the United States and Israel: Narrative Review and Qualitative Study (e50676)
Motti Haimi, Sheila Wheeler. 141

Evaluating the Usability of an mHealth App for Empowering Cancer Survivors With Disabilities: Heuristic Evaluation and Usability Testing (e51522)
Rachel Adler, Kevin Baez, Paulina Morales, Jocelyn Sotelo, David Victorson, Susan Magasi. 159

User Perceptions of Visual Clot in a High-Fidelity Simulation Study: Mixed Qualitative-Quantitative Study (e47991)
Greta Gasciauskaite, Clara Castellucci, Amos Malorgio, Alexandra Budowski, Giovanna Schweiger, Michaela Kolbe, Daniel Fries, Bastian Grande, Christoph Nöthiger, Donat Spahn, Tadzio Roche, David Tscholl, Samira Akbas. 173

Digital Care Pathway for Patients With Sleep Apnea in Specialized Care: Mixed Methods Study (e47809)
Jari Haverinen, Terttu Harju, Hanna Mikkonen, Pia Liljamo, Miia Turpeinen, Jarmo Reponen. 185

User Experience Evaluation of a Spinal Surgery Robot: Workload, Usability, and Satisfaction Study (e54425)
Hyeonkyeong Choi, Seunghee Kim, Wonseuk Jang. 204

Interactive Electronic Pegboard for Enhancing Manual Dexterity and Cognitive Abilities: Instrument Usability Study (e56357)
Shih-Ying Chien, Alice Wong, Ching-Yi Wu, Sara Beckman. 231

Effects of User Experience in Automated Information Processing on Perceived Usefulness of Digital Contact-Tracing Apps: Cross-Sectional Survey Study (e53940)
Tim Schrills, Lilian Kojan, Marthe Gruner, André Calero Valdez, Thomas Franke. 242

Health Care Professionals' Perspectives Before and After Use of eDialogue for Team-Based Digital Communication Across Settings: Qualitative Study (e53391)
Lili Jensen, Ole Rahbek, Rikke Lauritsen, Søren Kold, Birthe Dinesen. 259

Exploring the Role of Complexity in Health Care Technology Bottom-Up Innovations: Multiple-Case Study Using the Nonadoption, Abandonment, Scale-Up, Spread, and Sustainability Complexity Assessment Tool (e50889)
Ulla Hellstrand Tang, Frida Smith, Ulla Karilampi, Andreas Gremyr. 272

Provider Adoption of mHealth in Rural Patient Care: Web-Based Survey Study (e55443)
Bryan Weichelt, Rick Burke, Burney Kieke, Matt Pilz, Neel Shimpi. 288

Leveraging Generative AI Tools to Support the Development of Digital Solutions in Health Care Research: Case Study (e52885)
Danissa Rodriguez, Katharine Lawrence, Javier Gonzalez, Beatrix Brandfield-Harvey, Lynn Xu, Sumaiya Tasneem, Defne Levine, Devin Mann. 2 9 7

Testing Two Online Symptom Checkers With Vulnerable Groups: Usability Study to Improve Cognitive Accessibility of eHealth Services (e45275)
Kaisa Savolainen, Sari Kujala. 311

Head Protection Device for Individuals at Risk for Head Injury due to Ground-Level Falls: Single Trauma Center User Experience Investigation (e54854)
Susan Haag, John Kepros. 324

XSL•FO
RenderX

A Digital Health Intervention Platform (Active and Independent Management System) to Enhance the Rehabilitation Experience for Orthopedic Joint Replacement Patients: Usability Evaluation Study (e50430)
Petros Papadopoulos, Mario Soflano, Thomas Connolly. 332

Evaluation of the Parkinson's Remote Interactive Monitoring System in a Clinical Setting: Usability Study (e54145)
Bronwyn Bridges, Jake Taylor, John Weber. 347

Assessing the Relationship Between Digital Trail Making Test Performance and IT Task Performance: Empirical Study (e49992)
Tanguy Depauw, Jared Boasen, Pierre-Majorique Léger, Sylvain Sénécal. 366

Physical Therapists' Acceptance of a Wearable, Fabric-Based Sensor System (Motion Tape) for Use in Clinical Practice: Qualitative Focus Group Study (e55246)
Audrey Lee, Patricia Dionicio, Emilia Farcas, Job Godino, Kevin Patrick, Elijah Wyckoff, Kenneth Loh, Sara Gombatto. 379

Physicians' and Patients' Expectations From Digital Agents for Consultations: Interview Study Among Physicians and Patients (e49647)
Andri Färber, Christiane Schwabe, Philipp Stalder, Mateusz Dolata, Gerhard Schwabe. 394

Preferences for mHealth Intervention to Address Mental Health Challenges Among Men Who Have Sex With Men in Nepal: Qualitative Study (e56002)
Kamal Gautam, Camille Aguilar, Kiran Paudel, Manisha Dhakal, Jeffrey Wickersham, Bibhav Acharya, Sabitri Sapkota, Keshab Deuba, Roman Shrestha. 412

Patient Perspectives on Communication Pathways After Orthopedic Surgery and Discharge and Evaluation of Team-Based Digital Communication: Qualitative Exploratory Study (e49696)
Lili Jensen, Ole Rahbek, Rikke Lauritsen, Søren Kold, Birthe Dinesen. 424

The Asthma App as a New Way to Promote Responsible Short-Acting Beta2-Agonist Use in People With Asthma: Results of a Mixed Methods Pilot Study (e54386)
Liselot van den Berg, Cynthia Hallensleben, Lisa Vlug, Niels Chavannes, Anke Versluis. 441

A Smartphone App to Support Self-Management for People Living With Sjögren's Syndrome: Qualitative Co-Design Workshops (e54172)
Claire McCallum, Miglena Campbell, John Vines, Tim Rapley, Jason Ellis, Vincent Deary, Katie Hackett. 456

Effective Communication Supported by an App for Pregnant Women: Quantitative Longitudinal Study (e48218)
Lukas Kötting, Vinayak Anand-Kumar, Franziska Keller, Nils Henschel, Sonia Lippke. 471

Comparing Attitudes Toward Different Consent Mediums: Semistructured Qualitative Study (e53113)
Xengie Doan, Arianna Rossi, Marietjie Botes, Annika Selzer. 487

Knowledge, Skills, and Experience With Technology in Relation to Nutritional Intake and Physical Activity Among Older Adults at Risk of Falls: Semistructured Interview Study (e52575)
Julie Kikkenborg, Emma Magelund, Maria Riise, Lars Kayser, Rikke Terp. 503

Health Technology Access and Peer Support Among Digitally Engaged People Experiencing Homelessness: Qualitative Study (e55415)
Nóra Radó, Sándor Békási, Zsuzsa Gy rffy. 515

Digital Lifestyle Interventions for Young People With Mental Illness: A Qualitative Study Among Mental Health Care Professionals (e53406)
Chelsea Sawyer, Rebekah Carney, Lamiece Hassan, Sandra Bucci, John Sainsbury, Karina Lovell, John Torous, Joseph Firth. 530

XSL•FO
RenderX

Alarm Management in Intensive Care: Qualitative Triangulation Study (e55571)
Lina Mosch, Meltem Sümer, Anne Flint, Markus Feufel, Felix Balzer, Frauke Mörike, Akira-Sebastian Poncette. 555

Usability of an App for Medical History Taking in General Practice From the Patients' Perspective: Cross-Sectional Study (e47755)
Klara Albrink, Dominik Schröder, Carla Joos, Frank Müller, Eva Noack. 566

A Novel Continuous Real-Time Vital Signs Viewer for Intensive Care Units: Design and Evaluation Study (e46030)
Shiming Yang, Samuel Galvagno, Neeraj Badjatia, Deborah Stein, William Teeter, Thomas Scalea, Stacy Shackelford, Raymond Fang, Catriona Miller, Peter Hu, VS viewer study group. 581

Characterizing and Comparing Adverse Drug Events Documented in 2 Spontaneous Reporting Systems in the Lower Mainland of British Columbia, Canada: Retrospective Observational Study (e52495)
Erica Lau, Amber Cragg, Serena Small, Katherine Butcher, Corinne Hohl. 593

Usability and Evaluation of a Health Information System in the Emergency Department: Mixed Methods Study (e48445)
Christina Østervang, Charlotte Jensen, Elisabeth Coyne, Karin Dieperink, Annmarie Lassen. 609

Intention to Use an Electronic Community Health Information System Among Health Extension Workers in Rural Northwest Ethiopia: Cross-Sectional Study Using the Unified Theory of Acceptance and Use of Technology 2 Model (e47081)
Tesfahun Hailemariam, Asmamaw Atnafu, Lemma Gezie, Jens Kaasbøll, Jorn Klein, Binyam Tilahun. 623

Exploring the Use of Persuasive System Design Principles to Enhance Medication Incident Reporting and Learning Systems: Scoping Reviews and Persuasive Design Assessment (e41557)
Kiemute Oyibo, Paola Gonzalez, Sarah Ejaz, Tasneem Naheyan, Carla Beaton, Denis O'Donnell, James Barker. 637

Clinical Decision Support Requirements for Ventricular Tachycardia Diagnosis Within the Frameworks of Knowledge and Practice: Survey Study (e55802)
Zhao Hu, Min Wang, Si Zheng, Xiaowei Xu, Zhuxin Zhang, Qiaoyue Ge, Jiao Li, Yan Yao. 665

Usability of an Automated System for Real-Time Monitoring of Shared Decision-Making for Surgery: Mixed Methods Evaluation (e46698)
Christin Hoffmann, Kerry Avery, Rhiannon Macefield, Tadeáš Dvo ák, Val Snelgrove, Jane Blazeby, Della Hopkins, Shireen Hickey, Ben Gibbison, Leila Rooshenas, Adam Williams, Jonathan Aning, Hilary Bekker, Angus McNair, The ALPACA Study Team. 679

Influence of Disease-Related Stigma on Patients' Decisions to Upload Medical Reports to the German Electronic Health Record: Randomized Controlled Trial (e52625)
Niklas von Kalckreuth, Markus Feufel. 697

Academic Detailing as a Health Information Technology Implementation Method: Supporting the Design and Implementation of an Emergency Department–Based Clinical Decision Support Tool to Prevent Future Falls (e52592)
Hanna Barton, Apoorva Maru, Margaret Leaf, Daniel Hekman, Douglas Wiegmann, Manish Shah, Brian Patterson. 708

Electronic Immunization Registry in Rwanda: Qualitative Study of Health Worker Experiences (e53071)
Thaoussi Uwera, Mahima Venkateswaran, Kiran Bhutada, Eleni Papadopoulou, Enock Rukundo, David K Tumusiime, J Frøen. 719

The Solutions in Health Analytics for Rural Equity Across the Northwest (SHARE-NW) Dashboard for Health Equity in Rural Public Health: Usability Evaluation (e51666)
Elizabeth Heitkemper, Scott Hulse, Betty Bekemeier, Melinda Schultz, Greg Whitman, Anne Turner. 732

XSL•FO
RenderX

Implementing a Hospital Call Center Service for Mental Health in Uganda: User-Centered Design Approach (e53976)
Johnblack Kabukye, Rosemary Namagembe, Juliet Nakku, Vincent Kiberu, Marie Sjölinder, Susanne Nilsson, Caroline Wamala-Larsson. . . 7 4 4

Evaluation of a Computer-Aided Clinical Decision Support System for Point-of-Care Use in Low-Resource Primary Care Settings: Acceptability Evaluation Study (e47631)
Geletaw Tegenaw, Demisew Sori, Girum Teklemariam, Frank Verbeke, Jan Cornelis, Bart Jansen. 759

Assessing the Usability and Feasibility of Digital Assistant Tools for Direct Support Professionals: Participatory Design and Pilot-Testing (e51612)
Patrice Tremoulet, Andrea Lobo, Christina Simmons, Ganesh Baliga, Matthew Brady. 779

Assessing the Feasibility and Preliminary Effects of a Web-Based Self-Management Program for Chronic Noncancer Pain: Mixed Methods Study (e50747)
Pascale Marier-Deschenes, Anne Pinard, Laura Jalbert, Annie LeBlanc. 798

A Web-Based Peer Support Network to Help Care Partners of People With Serious Illness: Co-Design Study (e53194)
Elizabeth O'Donnell, Aricca Van Citters, Inas Khayal, Matthew Wilson, David Gustafson, Amber Barnato, Andrea Buccellato, Colleen Young, Megan Holthoff, Eugene Korsunskiy, Stephanie Tomlin, Amelia Cullinan, Alexandra Steinbaugh, Jennifer Hinson, Kristen Johnson, Andrew Williams, Ruth Thomson, Janet Haines, Anne Holmes, Ann Bradley, Eugene Nelson, Kathryn Kirkland. 818

Gamification Approach to Provide Support About the Deferral Experience in Blood Donation: Design and Feasibility Study (e50086)
Roberto Espinoza Chamorro, Luciano Santos, Yukiko Mori, Chang Liu, Goshiro Yamamoto, Tomohiro Kuroda. 832

Interest in mHealth Among Patients With Low Back Pain: Cross-Sectional Study (e48729)
Jonas Ipsen, Louise Sandal, Natalie Chang, Berit Schiøttz-Christensen, Karen Søgaard, Anders Hansen. 848

A Mobile Health App to Support Home-Based Aerobic Exercise in Neuromuscular Diseases: Usability Study (e49808)
Tim Veneman, Fieke Koopman, Sander Oorschot, Pien Koomen, Frans Nollet, Eric Voorn. 858

Validating the Effectiveness of the Patient-Centered Cancer Care Framework by Assessing the Impact of Work System Factors on Patient-Centered Care and Quality of Care: Interview Study With Newly Diagnosed Cancer Patients (e53053)
Safa Elkefi, Onur Asan. 880

Development of a Real-Time Dashboard for Overdose Touchpoints: User-Centered Design Approach (e57239)
Amey Salvi, Logan Gillenwater, Brandon Cockrum, Sarah Wiehe, Kaitlyn Christian, John Cayton, Timothy Bailey, Katherine Schwartz, Allyson Dir, Bradley Ray, Matthew Aalsma, Khairi Reda. 898

User-Centered Design and Usability of a Culturally Adapted Virtual Survivorship Care App for Chinese Canadian Prostate Cancer Survivors: Qualitative Descriptive Study (e49353)
Karen Young, Ting Xiong, Rachel Lee, Ananya Banerjee, Myles Leslie, Wellam Ko, Quynh Pham. 926

Effectiveness and User Perception of an In-Vehicle Voice Warning for Hypoglycemia: Development and Feasibility Trial (e42823)
Caterina Bérubé, Vera Lehmann, Martin Maritsch, Mathias Kraus, Stefan Feuerriegel, Felix Wortmann, Thomas Züger, Christoph Stettler, Elgar Fleisch, A Kocaballi, Tobias Kowatsch. 937

XSL•FO
RenderX

A Closed-Loop Falls Monitoring and Prevention App for Multiple Sclerosis Clinical Practice: Human-Centered Design of the Multiple Sclerosis Falls InsightTrack (e49331)
Valerie Block, Kanishka Koshal, Jaeleene Wijangco, Nicolette Miller, Narender Sara, Kyra Henderson, Jennifer Reihm, Arpita Gopal, Sonam Mohan, Jeffrey Gelfand, Chu-Yueh Guo, Lauren Oommen, Alyssa Nylander, James Rowson, Ethan Brown, Stephen Sanders, Katherine Rankin, Courtney Lyles, Ida Sim, Riley Bove. 953

Digital Triage Tools for Sexually Transmitted Infection Testing Compared With General Practitioners' Advice: Vignette-Based Qualitative Study With Interviews Among General Practitioners (e49221)
Kyma Schnoor, Anke Versluis, Niels Chavannes, Esther Talboom-Kamp. 971

Evaluating Users' Experiences of a Child Multimodal Wearable Device: Mixed Methods Approach (e49316)
Nancy McElwain, Meghan Fisher, Camille Nebeker, Jordan Bodway, Bashima Islam, Mark Hasegawa-Johnson. 981

Usability and User Experience of an mHealth App for Therapy Support of Patients With Breast Cancer: Mixed Methods Study Using Eye Tracking (e50926)
Carolin Anders, Preetha Moorthy, Laura Svensson, Julia Müller, Oliver Heinze, Petra Knaup, Markus Wallwiener, Thomas Deutsch, Thao-Vy Le, Lina Weinert. 997

Feasibility and Design Factors for Home-Based Pulmonary Rehabilitation of Patients With Chronic Obstructive Pulmonary Disease and Chronic Lung Diseases Based on a People-Object-Environment Framework: Qualitative Interview Study (e51150)
Shih-Ying Chien, Alice Wong, Winston Tseng, Han-Chung Hu, Hsiu-Ying Cho. 1014

A Technological Tool Aimed at Self-Care in Patients With Multimorbidity: Cross-Sectional Usability Study (e46811)
Rodrigo Medina-García, Juan López-Rodríguez, Cristina Lozano-Hernández, Verónica Ruiz Bejerano, Paride Criscio, Isabel Del Cura-González, TeNDER Atención Primaria Group. 1028

Multimodal In-Vehicle Hypoglycemia Warning for Drivers With Type 1 Diabetes: Design and Evaluation in Simulated and Real-World Driving (e46967)
Caterina Bérubé, Martin Maritsch, Vera Lehmann, Mathias Kraus, Stefan Feuerriegel, Thomas Züger, Felix Wortmann, Christoph Stettler, Elgar Fleisch, A Kocaballi, Tobias Kowatsch. 1043

Usability Comparison Among Healthy Participants of an Anthropomorphic Digital Human and a Text-Based Chatbot as a Responder to Questions on Mental Health: Randomized Controlled Trial (e54581)
Almira Thunström, Hanne Carlsen, Lilas Ali, Tomas Larson, Andreas Hellström, Steinn Steingrimsson. 1058

Identifying Factors of User Acceptance of a Drone-Based Medication Delivery: User-Centered Design Approach (e51587)
Franziska Fink, Ivonne Kalter, Jenny-Victoria Steindorff, Hans Helmbold, Denny Paulicke, Patrick Jahn. 1074

Facilitators of and Barriers to the Use of a Digital Self-Management Service for Diagnostic Testing: Focus Group Study With Potential Users (e45115)
Kyma Schnoor, Esther Talboom-Kamp, Muamer Hajti , Niels Chavannes, Anke Versluis. 1092

Evaluating the Energy Efficiency of Popular US Smartphone Health Care Apps: Comparative Analysis Study Toward Sustainable Health and Nutrition Apps Practices (e58311)
Abdullah Almasri, Tatyana El-Kour, Liliana Silva, Yousef Abdulfattah. 1105

Accessibility, Relevance, and Impact of a Symptom Monitoring Tool for Home Hospice Care: Theory Elaboration and Qualitative Assessment (e51789)
Karla Washington, Debra Oliver, Allison Donehower, Patrick White, Jacquelyn Benson, Patrick Lyons, George Demiris. 1116

XSL•FO
RenderX

Usability, Acceptability, and Preliminary Effectiveness of a Peer-Delivered and Technology-Supported Mental Health Intervention for Family Caregivers of People With Dementia: Field Usability Study (e41202)
Caroline Collins-Pisano, Amanda Leggett, David Gambee, Karen Fortuna. 1131

Smartphone App Designed to Collect Health Information in Older Adults: Usability Study (e56653)
Joanne Murabito, Jamie Faro, Yuankai Zhang, Angelo DeMalia, Alexander Hamel, Nakesha Agyapong, Hongshan Liu, Eric Schramm, David McManus, Belinda Borrelli. 1150

The Influence of Incentive-Based Mobile Fitness Apps on Users' Continuance Intention With Gender Moderation Effects: Quantitative and Qualitative Study (e50957)
Aaya Faizah, Alifah Hardian, Rania Nandini, Putu Handayani, Nabila Harahap. 1163

Using a Smartwatch App to Understand Young Adult Substance Use: Mixed Methods Feasibility Study (e50795)
Sahiti Kunchay, Ashley Linden-Carmichael, Saeed Abdullah. 1179

Perspectives on the COVID-19 Vaccination Rollout in 17 Countries: Reflexive Thematic and Frequency Analysis Based on the Strengths, Weaknesses, Opportunities, and Threats (SWOT) Framework (e44258)
Vanja Kopilaš, Khrystyna Nasadiuk, Lucia Martinelli, Lenka Lhotska, Zoran Todorovic, Matjaz Vidmar, Helena Machado, Anna Svalastog, Sre ko Gajovi 1196

The Role of Social Media in the Experiences of COVID-19 Among Long-Hauler Women: Qualitative Study (e50443)
Camryn Garrett, Atefeh Aghaei, Abhishek Aggarwal, Shan Qiao. 1211

Usability and Acceptability of a Conversational Agent Health Education App (Nthabi) for Young Women in Lesotho: Quantitative Study (e52048)
Elizabeth Nkabane-Nkholongo, Mathildah Mpata-Mokgatle, Brian Jack, Clevanne Julce, Timothy Bickmore. 1225

Stimulating Preconception Care Uptake by Women With a Vulnerable Health Status Through a Mobile Health App (Pregnant Faster): Pilot Feasibility Study (e53614)
Sharissa Smith, Babette Bais, Hafez Ismaili M'hamdi, Maartje Schermer, Régine Steegers-Theunissen. 1238

Efficacy of the QuitSure App for Smoking Cessation in Adult Smokers: Cross-Sectional Web Survey (e49519)
Gregory Goldgof, Shweta Mishra, Kriti Bajaj. 1254

Digital Adherence Technologies Linked to Mobile Money Incentives for Medication Adherence Among People Living With Tuberculosis: Mixed Methods Feasibility and Acceptability Study (e47996)
Angella Musiimenta, Wilson Tumuhimbise, Esther Atukunda, Aaron Mugaba, Sebastian Linnemayr, Jessica Haberer. 1267

A Multicomponent Intervention (POSSIBLE) to Improve Perceived Risk for HIV Among Black Sexual Minority Men: Feasibility and Preliminary Effectiveness Pilot Study (e54739)
Derek Dangerfield II, Janeane Anderson, Charleen Wylie, Ricky Bluthenthal, Chris Beyrer, Jason Farley. 1283

Characteristics, Opportunities, and Challenges of Osteopathy Based on the Perceptions of Osteopaths in Austria: Qualitative Interview Study (e45302)
Jonas Manschel, Jan Porthun, Ulrike Winkler, Jean Beuckels, David Martin. 1293

A Digital Behavioral Activation Intervention (JuNEX) for Pregnant Women With Subclinical Depression Symptoms: Explorative Co-Design Study (e50098)
Elisa Mancinelli, Silvia Gabrielli, Silvia Salcuni. 1307

Trust in and Acceptance of Artificial Intelligence Applications in Medicine: Mixed Methods Study (e47031)
Daria Shevtsova, Anam Ahmed, Iris Boot, Carmen Sanges, Michael Hudecek, John Jacobs, Simon Hort, Hubertus Vrijhoef. 1325

XSL•FO
RenderX

A Machine Learning Approach with Human-AI Collaboration for Automated Classification of Patient Safety Event Reports: Algorithm Development and Validation Study (e53378)
Hongbo Chen, Eldan Cohen, Dulaney Wilson, Myrtede Alfred. 1342

Testing the Feasibility and Acceptability of Using an Artificial Intelligence Chatbot to Promote HIV Testing and Pre-Exposure Prophylaxis in Malaysia: Mixed Methods Study (e52055)
Min Cheah, Yan Gan, Frederick Altice, Jeffrey Wickersham, Roman Shrestha, Nur Salleh, Kee Ng, Iskandar Azwa, Vimala Balakrishnan, Adeeba Kamarulzaman, Zhao Ni. 1355

The Impact of Performance Expectancy, Workload, Risk, and Satisfaction on Trust in ChatGPT: Cross-Sectional Survey Analysis (e55399)
Avishek Choudhury, Hamid Shamszare. 1375

The Role of Coherent Robot Behavior and Embodiment in Emotion Perception and Recognition During Human-Robot Interaction: Experimental Study (e45494)
Laura Fiorini, Grazia D'Onofrio, Alessandra Sorrentino, Federica Cornacchia Loizzo, Sergio Russo, Filomena Ciccone, Francesco Giuliani, Daniele Sancarlo, Filippo Cavallo. 1386

The Effect of a Video-Assisted Health Education Program Followed by Peer Education on the Health Literacy of COVID-19 and Other Infectious Diseases Among School Children: Quasi-Randomized Controlled Trial (e43943)
Xiaojuan Zhang, Yingkun Wen, Ning Han, Yawen Jiang. 1400

Barriers to and Facilitators of Key Stakeholders Influencing Successful Digital Implementation of Remote Monitoring Solutions: Mixed Methods Analysis (e49769)
Fahad Iqbal, Ravi Aggarwal, Meera Joshi, Dominic King, Guy Martin, Sadia Khan, Mike Wright, Hutan Ashrafian, Ara Darzi. 1411

Short Papers

Assessing Differences in mHealth Usability and App Experiences Among Young African American Women: Secondary Analysis of a Randomized Controlled Trial (e51518)
Claudia Opper, Felicia Browne, Brittni Howard, William Zule, Wendee Wechsberg. 225

Designing a Novel Digitally Delivered Antiracism Intervention for Mental Health Clinicians: Exploratory Analysis of Acceptability (e52561)
Tashalee Brown, Habiba Amir, Drew Hirsch, Madeline Jansen. 873

Viewpoints

Implementation of Anxiety UK's Ask Anxia Chatbot Service: Lessons Learned (e53897)
Luke Collins, Niamh Nicholson, Nicky Lidbetter, Dave Smithson, Paul Baker. 544

The Temperature Feature of ChatGPT: Modifying Creativity for Clinical Research (e53559)
Joshua Davis, Liesbet Van Bulck, Brigitte Durieux, Charlotta Lindvall. 1371

XSL•FO
RenderX

Original Paper

Acceptance and User Experiences of a Wearable Device for the Management of Hospitalized Patients in COVID-19–Designated Wards in Ho Chi Minh City, Vietnam: Action Learning Project

An Phuoc Luu[1], MPH; Truong Thanh Nguyen[2], MD; Van Thi Cam Cao[2], RN; Trinh Hoang Diem Ha[2], RN; Lien Thi Thu Chung[2], RN; Trung Ngoc Truong[2], MD; Tung Nguyen Le Nhu[2], MD; Khoa Bach Dao[2], MD; Hao Van Nguyen[2], MD; Phan Nguyen Quoc Khanh[1], MD; Khanh Thuy Thuy Le[1], MD; Luu Hoai Bao Tran[1], MD; Phung Tran Huy Nhat[1], BSc; Duc Minh Tran[1], PMD; Yen Minh Lam[1], MD; Catherine Louise Thwaites[1,3], BSc, MBBS, MD; Jacob Mcknight[3], BEng, MSc, DPhil; Nguyen Van Vinh Chau[1,2], MD; Jennifer Ilo Van Nuil[1,3], BA, MA, PhD; Vietnam ICU Translational Applications Laboratory (VITAL)[4]

[1]Oxford University Clinical Research Unit, Ho Chi Minh City, Vietnam

[2]Hospital for Tropical Diseases, Ho Chi Minh City, Vietnam

[3]Centre for Tropical Medicine and Global Health, Nuffleld Department of Medicine, University of Oxford, Oxford, United Kingdom

[4]See Acknowledgments

Corresponding Author:
Jennifer Ilo Van Nuil, BA, MA, PhD
Oxford University Clinical Research Unit
764 Vo Van Kiet
Ho Chi Minh City, 70000
Vietnam
Phone: 84 362620124
Email: jvannuil@oucru.org

Abstract

Background: Wearable devices have been used extensively both inside and outside of the hospital setting. During the COVID-19 pandemic, in some contexts, there was an increased need to remotely monitor pulse and saturated oxygen for patients due to the lack of staff and bedside monitors.

Objective: A prototype of a remote monitoring system using wearable pulse oximeter devices was implemented at the Hospital for Tropical Diseases in Ho Chi Minh City, Vietnam, from August to December 2021. The aim of this work was to support the ongoing implementation of the remote monitoring system.

Methods: We used an action learning approach with rapid pragmatic methods, including informal discussions and observations as well as a feedback survey form designed based on the technology acceptance model to assess the use and acceptability of the system. Based on these results, we facilitated a meeting using user-centered design principles to explore user needs and ideas about its development in more detail.

Results: In total, 21 users filled in the feedback form. The mean technology acceptance model scores ranged from 3.5 (for perceived ease of use) to 4.4 (for attitude) with behavioral intention (3.8) and perceived usefulness (4.2) scoring in between. Those working as nurses scored higher on perceived usefulness, attitude, and behavioral intention than did physicians. Based on informal discussions, we realized there was a mismatch between how we (ie, the research team) and the ward teams perceived the use and wider purpose of the technology.

Conclusions: Designing and implementing the devices to be more nurse-centric from their introduction could have helped to increase their efficiency and use during the complex pandemic period.

(JMIR Hum Factors 2024;11:e44619) doi:10.2196/44619

KEYWORDS

vital signs; wearable devices; action learning; technology acceptance model; TAM; COVID-19; user-centered design; wearables; remote monitoring; technology acceptance; oximeter

XSL•FO
RenderX

Introduction

The popularity of portable wearable technologies that monitor health has increased substantially over the past decade due to their perceived utility, relatively simple implementation, and immediate feedback [1]. Wearable technology is used in both personal and clinical settings, and more recently in the context of the COVID-19 pandemic for diagnosis, remote monitoring, and other applications in both inpatient and outpatient settings [1-3]. Using wearable devices for COVID-19 care can result in infection control by reducing the amount of time that health care workers (HCWs) are physically with patients and providing continuous monitoring of vital signs for the early identification and potential treatment of deteriorating patients [2]. Specifically, remote monitoring of oxygen saturation using wearable devices became increasingly common during COVID-19 in hospital settings [4,5].

Despite the potential benefits, there have been many challenges noted in implementing and using wearable devices during COVID-19, including technical, social, and political spheres [1]. Technical challenges often include battery life, Wi-Fi or Bluetooth connections, and device communication. A few examples of social challenges are patients lacking technological confidence (eg, in older patients) and repeated device alerts or continuous monitoring making patients nervous, while political challenges could include regulatory issues for expanding the approval of devices for COVID-19–related medical situations [1]. Many of these challenges may be enhanced in low- and middle-income countries (LMICs), while the need for such integration is crucial, especially during pandemic situations [6,7].

There are several studies exploring the technical challenges of integrating wearable devices in trial settings during COVID-19 (eg, see [8]), but there is a lack of research surrounding the acceptability of such devices within these contexts and how attitudes may impact actual use [9]. Portable wearable devices could be a potential solution to allow for continuous monitoring of vital signs remotely and affordably for COVID-19 wards in LMIC settings; however, while advantageous, these devices cannot meet their full potential if the users do not agree to use them or realize their potential value [10]. Understanding user perceptions and needs as well as the context in which the technology is implemented is crucial for successful implementation [1]. User-centered approaches stress the importance of integrating both human factors and technical factors [11] while also paying attention to avoid excluding certain populations in the design [7]. User-centered approaches have been cited as a “critical success factor” in a variety of health-related technology projects [12].

From August to December 2021, when COVID-19 cases were increasing more rapidly than at any time previously in Ho Chi Minh City, Vietnam, there was an opportunity to integrate a prototype wearable device and monitoring system into the COVID-19–designated wards at the Hospital for Tropical Diseases (HTD). At this time, the HTD was overwhelmed with patients with COVID-19 and we needed to deploy something urgently that could help. Using pragmatic methods during the rollout of the device, we describe stakeholders’ use of the wearable device, aspects of acceptability, and under which circumstances its use would be most beneficial for improving the care of patients with COVID-19. The primary objective of this work was to support the implementation process of the wearable device in the hospital to improve patient care during a catastrophic period of the COVID-19 pandemic in Ho Chi Minh City, Vietnam.

Methods

Study Setting

This work took place within a larger project called the Vietnam ICU Translation Application Laboratory (VITAL) at the Oxford University Clinical Research Unit (OUCRU) and HTD. The goal of VITAL is to design and implement innovative technologies to improve patient care within the intensive care unit (ICU) at the HTD, with a longer-term goal of expanding these technologies regionally. In addition to the clinical and technological studies, there is an ethnographic study to explore the sociotechnical contexts of the ICU at the HTD and within ICUs in Vietnam more broadly. The VITAL multidisciplinary team was in place at the start of the COVID-19 pandemic.

In the first 100 days of the COVID-19 pandemic, Vietnam rapidly implemented a variety of public health measures resulting in relatively few cases and zero deaths [13]. Since that time, there were a few concentrated outbreaks (for example, in Da Nang in July 2020 and December 2020 in northern Vietnam). In May 2021, the cases started to increase on a countrywide basis, and by August 2021, the hospitals began to fill with patients with COVID-19. It was within this pandemic context that the wearable device was implemented at the HTD, and the VITAL study teams worked together throughout to improve its implementation.

The wearable device was selected by the company and was already integrated into a locally developed platform based on an available application programming interface, licensing, and availability. The device was medical grade and measured heart rate and blood oxygen levels, similar to a pulse oximeter. The wearable device was battery powered and each one connected to a tablet that was kept at the patients’ bedside. The tablets had a 3G or 4G sim card and sent the data to a cloud where multiple patients’ data were viewable by HCWs outside the patients’ rooms and isolation area.

Study Design

The aim of this work was to support the ongoing implementation of the wearable device rather than to follow a predefined, replicable study protocol, as would be used in trial settings, for example. Therefore, the work here describes the pragmatic rollout of the device. We used an action learning approach, including integration of multiple methods to assess the use and acceptability of the wearable device [14]. Action learning approaches rely on an iterative process of assessing local contexts, learning from relevant stakeholders, and using the information to improve an implementation or further develop a technology specific to the context [15,16]. As the wearable device started to be implemented in the HTD wards, our team

XSL•FO
RenderX

of HCWs, social scientists, and technology developers took the opportunity to work together to inform the implementation. Therefore, we adapted the methods as the situation changed and more insights were gained [14].

Participants

Potential participants included the HCWs from the HTD who were using the device in the wards during the implementation and corresponding ward heads. We estimated that a total of 30 doctors and 60 nurses would have worked in the wards where the wearable device was implemented and potentially used it in some form; therefore, we planned to recruit participants from this larger sample.

Data Collection Methods

Informal Discussions and Observations

We used an iterative process of engaging in informal discussions coupled with sense-checking discussions and observations during the implementation period. The informal and sense-checking discussions and observations were conducted with the team who was working directly in the wards, as well as with head nurses from the wards where the wearable device was being implemented. The informal discussions and observations were conducted during the implementation of the device.

Feedback Survey Form

We created the feedback form based on the components of the technology acceptance model (TAM) to assess the use and acceptance of the device. The TAM is used in a variety of disciplines to determine how individuals accept (or not) and use (or not) a given technology. Davis [17] developed this model based on components from the theory of reasoned action [18] and it consists of the following variables: use motivation (with perceived ease of use and perceived usefulness) and behavioral intention [17,19]. The model suggests that an individual will accept the use of a technology (ie, their behavioral intention) based on their perception of the technology's usefulness and ease of use. Perceived usefulness refers to the perception that using the technology will enhance one's work; for example, the wearable device will provide physicians and nurses some advantages (eg, remote monitoring). Perceived ease of use refers to the perception that the use does not add more work or effort to the work that could be enhanced; for example, using the wearable device will not increase nurses' workload, despite its utility and simplicity [17]. The TAM framework was expanded twice to include attitudes as well as several other external factors [20]. The use of the TAM in health research has shown how perceived usefulness and perceived ease of use relates positively to attitude and behavioral intention [21]. The TAM has been criticized for being insensitive to the context or social factors, being simplistic, and following an assumption that users are rational decision makers, when indeed other factors play into decision making [22-24]. We used the TAM framework for its simplicity and because the categories of perceived usefulness and perceived ease of use were of relevance, but we also integrated other data collection methods alongside it to counter these limitations to some extent.

Based on the components of the TAM, we included 23 questions related to usefulness (n=5), ease of use (n=5), attitude (n=5), and behavioral intent (n=8) [25]. We asked these questions using a 5-point Likert scale (with scores of 5 being more favorable). We also added 2 open-ended questions and collected a variety of relevant demographic information (Multimedia Appendix 1). We piloted the tool in both English and Vietnamese and adjusted the form as needed. We used Google forms for electronic self-completion of the form and offered paper forms for hand-written self-completion. We explained the feedback form to the ward staff during team meetings and provided the link. The feedback form was distributed and completed in Vietnamese. We kept the feedback form link open for 7 weeks in total and started data collection after the implementation had been integrated into the wards so that users would have had experience using the device.

User-Centered Workshop

We held a user-centered workshop with a selection of HTD ward staff to explore user needs and ideas for development in more detail. Because we already had the technology and knew the spaces where implementation would be held, we followed an adapted version of the process described by Cooper et al [26]. With this approach, the workshop participants and facilitators set the scene as a busy COVID-19 ward during the peak of the pandemic. Then, the facilitators described the shells of users (personas), including a nurse and a doctor persona shell, and we had the workshop attendees describe who they imagined the nurse and doctor to be, as well as their behaviors and needs and the values each user group would find most essential. We based the conversation on the wearable technology that the participants had already used. Then, the group discussed solutions to the issues identified [26].

Data Analysis

Using the principles of action learning, we integrated the responses from informal discussions and observations into subsequent data collection, as well as summarized the content and grouped it into themes. For the analysis of the feedback survey form, we calculated mean scores for each variable and compared scores by profession. For the open-ended survey questions, we used content coding to summarize the responses topically. We presented the demographic data descriptively. We documented the responses from the user-centered design workshop as notes and summarized the results into main themes.

Ethical Considerations

In this paper, we are describing the processes that occurred as part of the development and implementation of a monitoring system; therefore, the work did not require ethics approval. Prior to the initiation of any activities, we held a meeting with ward heads to describe the work in more detail and obtain their agreement.

Results

Device Implementation Within the HTD Context

The wearable device was implemented in 3 wards starting in August 2021, including the adult ICU, Ward A, and Ward E.

XSL•FO
RenderX

We describe the implementation over a 5-month period from August to December 2021. During this period, these wards changed from COVID-19–designated and then back again to routine patient care settings, depending on the number of patients. Although the HTD was one of the COVID-19–designated hospitals, throughout the pandemic they offered routine patient care for specific diseases (eg, tetanus).

In addition to the rapidly changing physical spaces, the hospital management quickly deployed remote monitoring capacity using existing closed-circuit television cameras as a temporary solution to monitor very sick patients from outside the patients' rooms. The remote monitoring was useful as it allowed for multitasking and prevented nurses and doctors from checking on patients more routinely in person. The hospital wards were at capacity during the study period. Prior to the pandemic, however, it was not unusual for the wards at the HTD to often be at maximum patient capacity. For example, in the adult ICU or during the rainy season, the number of dengue patients increases dramatically and the wards tend to be full.

Also, the workflow was organized differently during the pandemic period. Instead of nurses taking care of a few specific patients for the whole shift, 2 nurses and 1 doctor would instead go into the ward (in full personal protective equipment) as a team for 3 hours at a time while the other 2 nurses on shift completed admin work in the office. This meant that more coordination was needed, and often the team with the patients "need[ed] someone else to be [their] memory" as it was not easy to remember everything about all patients. The health care team's workload, especially that of the nurses, ended up being more extensive for many reasons. One important reason is that, because of COVID-19 restrictions, there were also no families allowed in the wards who would help to look after patients in non–COVID-19 times; therefore, the majority of the care was left to the nurses. The patients were also more severely ill than previously in these wards and required more care by fewer staff.

Device Use and Acceptability

When we first distributed the feedback form, out of 90 potential participants, only 22 completed the survey (19 electronic and 3 paper forms), and 1 person stated that they did not use the technology and therefore no responses were recorded for that participant. Of the 21 respondents who completed the feedback form, 48% (n=10) were doctors and 48% (n=10) were nurses, with 52% (n=11) of the participants coming from Ward E (Table 1).

Overall, when assessing the TAM variables, the mean (SD) scores ranged from 3.6 (0.8) for perceived ease of use to 4.4 (0.6) for attitude, with behavioral intention (mean 3.9, SD 0.6) and perceived usefulness (mean 4.2, SD 0.7) scoring in between. Those working as nurses scored higher on perceived usefulness, perceived ease of use, attitude, and behavioral intention than did physicians (Table 2).

When asked, as an open-ended question, why participants would or would not use the wearable device in the future, of the 19 responses inputted, 15 participants wrote that they would use the system because of its convenience and usefulness in monitoring patients. However, in 2 of those responses, they also added comments that the device had limited perceived accuracy and transmission problems. Of the remaining 4 participants, 1 participant simply stated that the monitor was still in use, 2 participants wrote that they did not use the system anymore due to job location changes, and 1 participant wrote a few sentences about why the wearable device is not the "best choice," highlighting its limited battery life, how the system had become additional work for the already overworked staff, and how it is not yet completely implemented.

Table 1. Demographic characteristics of the survey respondents (n=21).

Characteristic	Value
Gender, n (%)	
Women	13 (62)
Men	8 (38)
Age (years), median (IQR)	35 (30-38)
Occupation, n (%)	
Doctor	10 (48)
Nurse	10 (48)
Other: nurses' aid	1 (5)
Primary ward during the implementation phase, n (%)	
Adult intensive care unit	6 (29)
Ward A	1 (5)
Ward D	3 (14)
Ward E	11 (52)

XSL•FO
RenderX

Table 2. Mean technology acceptance model (TAM) scores by variable. The maximum score was 5.

TAM variable	All participants, mean (SD)	Nurses, mean (SD)	Doctors, mean (SD)
Perceived usefulness	4.2 (0.7)	4.3 (0.8)	4.1 (0.6)
Perceived ease of use	3.6 (0.8)	3.8 (0.7)	3.4 (0.8)
Attitude	4.4 (0.6)	4.6 (0.6)	4.2 (0.6)
Behavioral intention	3.9 (0.6)	4.0 (0.6)	3.7 (0.7)

Integrating User Perceptions for Improved Implementation

As part of the action learning process, we supplemented the feedback form results with data from the observations and informal discussions during the 5-month period. There were 3 main observations. First, there was a mismatch between how we (ie, the research team) and the ward teams perceived the use of the technology. We quickly realized, from our observations and from informal discussions with the implementation team, that many of the nurses either did not use the wearable device or did not think that they used it even if they used it in some aspect (eg, connecting the device for the patients or changing batteries). Even after we clarified what we meant by "use," there were still not additional participants who filled in the feedback form because they felt like they did not use the technology.

Second, the ward teams had varying perceptions of the technologies that are routinely implemented by the OUCRU team in the HTD wards as part of research projects. We heard from informal discussions with colleagues that the nurses assumed the wearable devices were from a research project, as is often the case with OUCRU projects, and therefore the nurses, in particular, ignored the device even if they had some role in its use. They did not see its potential benefit.

Finally, in order to make the device more useful for the ward staff, we realized during the meetings and informal discussions with the team that we needed to make the implementation and use of the device more "pro-nurse," meaning we would need to emphasize how the device and its data were also useful and relevant to them. When discussing with the head nurse, the data were only displayed on the main screen in the staff room for one department. One suggestion was to move the tablet to the wall so that the nurses and others in the room (including the patients) could potentially see their vital signs. Because the devices and corresponding data were not in sight, it was easy to think that it was not relevant for the nurses and made it easier for them to ignore the device while with the patients.

User-Centered Design Workshop

With the information we had learned from the informal discussions, observations, and feedback form, we held a follow-up workshop on January 17, 2022, to discuss how we could make better use of the technology in the wards in COVID-19 situations in the future. The attendees included 2 doctors (1 man and 1 woman) and 3 nurses (2 women and 1 man). The participants discussed the behaviors and needs of the nurse and doctor persona. For both roles, the needs centered on having equipment and improved coordination. The nurses also mentioned more training needs, while the doctors' needs were about the accuracy of monitoring (Textboxes 1 and 2).

There were 3 main value prop themes, including medical, technical, and patient themes. For medical aspects, the attendees discussed how the device should be able to provide highly accurate data, with appropriate alarms and cut-offs. For the technical theme, the device and software should be simple to connect and use, with a long battery life and stable connections during charging or switching devices. The display should be large and clear, and the data should be stored for a long period of time (ie, 7-10 days). Finally, for the patient theme, the device should be comfortable for the patients to wear to avoid them removing it.

There were several solutions discussed in the group to improve the use and efficacy of the wearable device (Table 3). Solutions included improving the credibility of the data, ideas to improve the ease of use, ways to make the alarms more consistent, and ideas for more ideal placement. One very specific issue that the group mentioned was that the alarms went off too much on the large display and the alarms were always red or black and blinking, and it was difficult to know if the device was turned off (due to patient discharge) or actually disconnected, which would require an intervention. The solution was to refresh the devices; however, if the alarms were excessive and not always indicating a real issue, trust in the device would remain low, so this was an important priority. They also suggested that the alarms and display on the tablet should be the same as the big screen, as they preferred screen consistency.

Another in-depth discussion was about moving the tablets to the walls and having the device plugged in all the time, which would solve the battery issues. They felt that the tablet could be set up on the wall but that brought up other issues about how to keep the device and watch safe after use. For some of the topics, the group used features of another wearable device that they had used in the wards in the past to inform their solutions (eg, device graphs and a line on the device for finger placement).

XSL•FO
RenderX

Textbox 1. Behaviors and needs of the nurses.

Participant: Nurse Van is a 36-year-old woman. She is an administrative nurse and has a management job. She likes to have fun and has a family and 2 children. She is also responsible for bringing the kids to school and back.

Behaviors:

- Visit and provide direct patient care and monitor vital signs
- Carry out medical orders (ie, medications, blood tests, and nutrition)
- Assess, monitor, and hand over patients
- Work night duty
- Night shifts inform doctors on vital signs as prescribed

Needs:

- Equipment (eg, to measure blood pressure, temperature, oxygen levels, and heart rate)
- Training on diseases
- Teamwork and coordination

Textbox 2. Behaviors and needs of the doctors.

Participant: Doctor Huong is a 30-year-old woman. She is flexible and very active. She is not married and has no children and currently lives in a hotel. She is on night shift every 4 nights, and at times she visits her home in another town in Ho Chi Minh City, which is far from the Hospital for Tropical Diseases.

Behaviors:

- Prescribe medications
- Update medical records
- Perform examinations and change treatments
- Data entry
- Check vital signs in patient rooms (with a portable monitor that they move around) for examination and to detect abnormalities

Needs:

- Equipment (eg, monitors)
- Coordination with nurses (progress: medical records)
- Re-evaluation and working with other doctors
- Accuracy of vital sign monitoring

Table 3. Solutions for improvement.

Topic	Specific solution
Data credibility	Adding a graph for signal strength
Ease of use	• Adding a finger placement mark on the device • Increasing the font size on the watch and tablet • Tablets should be fixed on the wall
Alarms	• Reduce the alarm colors and blinking on the screen • Use the same display on the screen and the tablets for consistency • Refresh the tablets for more accurate alarms
Battery issues	Keep the tablet plugged in
Device placement	Placement on the wall (but only with an increase in font size)

XSL•FO
RenderX

Discussion

The HTD and OUCRU teams, along with the technology company, rolled out the wearable device in an extremely complex pandemic situation with a prototype system. In the end, the team used the device on over 100 patients. We assessed the usability and acceptance of the device over the implementation period when COVID-19 cases were peaking in the hospital and into the period when the COVID-19 cases were reducing. Similar to the literature on the topics, we found that the importance of understanding the users and their experiences using the device was crucial to get the most use out of these technologies.

There was a mismatch between our perception of who was using the device and those who thought that they were using or benefiting from the device on the ground. From the start, the device was designed and set up with doctors in mind, but in practice, the nurses' roles and use were overlooked, even though they could also routinely use and benefit from the device. In our study, we found that the nurses who filled in the feedback form, on average, had slightly higher scores on 3 of the 4 TAM domains (ie, perceived usefulness, attitude, and behavioral intention), while the doctors, on average, scored the perceived ease of use slightly higher than the nurses. We know from the challenges with acquiring feedback that many nurses did not feel that they used the device even though they had some role in the device set-up and monitoring. Designing the device to be more nurse-centric from the early phases could have helped to increase the efficiency and definition of who is meant to use it. In the future, it is important to consider that the way the device is used might be dependent on the form of its use (eg, for triage, use in a pandemic emergency, or routine hospital use). We recommend the involvement of staff who could benefit from the technology, especially nurses in the hospital context, in the full implementation process. This could help to avoid mismatches in the perceptions of who the users are and who could and should benefit from the new technology. Research on integrating wearable devices during COVID-19 in Singapore also highlighted that device simplicity would encourage its use and the importance of making the technology fit into the current environment while not increasing or disrupting workflows [27].

The trust in the device and its data was an issue brought up several times during the implementation and feedback sessions. There are a variety of potential explanations for inconsistent data (eg, incorrect device placement or averaging of data); however, it reduced the credibility of the device for both doctors and, importantly, nurses. Data concerns about technology in clinic settings has been noted in other studies. For example, Faria et al [28] found that study clinicians reported that 36% of the data from a remote monitoring project were "invalid" for a variety of reasons, including low literacy of the patients and complexity of the device. Involvement from users from the beginning of the design and implementation process is crucial for design purposes but also to build trust and confidence in the devices [11]. While this project took place during COVID-19, which is a very specific circumstance, the broader findings resonate with research conducted prior to COVID-19 that focused on the implementation and scaling up of digital health technologies in LMICs. The recommendations also included integration of end-user feedback and engagement with all stakeholders throughout the design and implementation process [12]

There are limitations to this work. First, we did not collect data on the clinical worth or the accuracy of the data transmitted from the devices. Second, we focused on feedback from only heath care staff (ie, doctors and nurses), and from only a subset of those who perceived that they used the device, which may have excluded some users and limited the overall sample size. We did not include patients who could also inform device acceptance, especially if used in noncritical cases where patients are moving around and conscious. Finally, the implementation setting for this work is not typical of other hospital settings in Vietnam or possibly other LMICs, as the HTD is a large referral hospital with an international research institute attached to it.

In anticipation of future (novel) pandemic situations or integration of wearable technologies into a range of clinical settings more broadly, it is important to fully understand if and how the wearable devices could be used more effectively by doctors, and importantly, nurses in the wards, for monitoring of deteriorating patients, especially in LMICs where resources are already stretched. Using an action learning approach during the implementation process highlights the importance of integrating user perspectives, ideas, and solutions into development and design.

Acknowledgments

We would like to acknowledge Nguy n Đ i Phú and Nguy n Th Tin from the Hospital for Tropical Diseases and Nguy n Th Di m Trinh and Đ ng Ph ng Th o from the Oxford University Clinical Research Unit for their support and commitment in operating the remote monitoring system.

The members of the Vietnam ICU Translational Applications Laboratory (VITAL) group are as follows: Oxford University Clinical Research Unit, Ho Chi Minh City, Vietnam: An Phuoc Luu, Chanh Quang Ho, Duc Hong Du, Duc Minh Tran, Dung Thi Phuong Nguyen, Giang Thi Nguyen, Hai Bich Ho, Hien Van Ho, Hung Manh Trinh, Huy Quang Nguyen, Khanh Nguyen Quoc Phan, Khoa Dinh Van Le, Kien Trung Dang, Lam Khanh Phung, Lieu Thi Pham, Ngoc Thanh Nguyen, Nhat Tran Huy Phung, Phuong Thanh Le, Quyen Than Ha Nguyen, Thanh Thi Le Nguyen, Thy Bui Xuan Doan, Trieu Trung Huynh, Trinh Huu Khanh Dong, Van Minh Tu Hoang, Van Thi Thanh Ninh, Vuong Lam Nguyen, Yen Minh Lam, Sayem Ahmed, Joseph Donovan, Ronald Geskus, Evelyne Kestelyn, Angela Mcbride, Guy Thwaites, Louise Thwaites, Hugo Turner, Jennifer Ilo Van Nuil, and Sophie Yacoub; Hospital for Tropical Diseases, Ho Chi Minh City, Vietnam: Tam Thi Cao, Thuy Bich Duong, Duong Thi Hai

XSL•FO
RenderX

Ha, Nghia Dang Trung Ha, Chau Buu Le, Thu Ngoc Minh Le, Thao Thi Mai Le, Tai Thi Hue Luong, Phu Hoan Nguyen, Viet Quoc Nguyen, Nguyen Thanh Nguyen, Phong Thanh Nguyen, Anh Thi Kim Nguyen, Hao Van Nguyen, Duoc Van Thanh Nguyen, Chau Van Vinh Nguyen, Oanh Kieu Nguyet Pham, Van Thi Hong Phan, Qui Tu Phan, Tho Vinh Phan, and Thao Thi Phuong Truong; University of Oxford, Oxford, United Kingdom: David Clifton, Mike English, Shadi Ghiasi, Heloise Greeff, Jannis Hagenah, Ping Lu, Jacob McKnight, Chris Paton, and Tingting Zhu; Imperial College London, London, United Kingdom: Pantellis Georgiou, Bernard Hernandez Perez, Kerri Hill-Cawthorne, Alison Holmes, Stefan Karolcik, Damien Ming, Nicolas Moser, and Jesus Rodriguez Manzano; King's College London, London, United Kingdom: Alberto Gomez, Hamideh Kerdegari, Marc Modat, and Reza Razavi; ETH Zurich, Zurich, Switzerland: Abhilash Guru Dutt, Walter Karlen, Michaela Verling, and Elias Wicki; The University of Melbourne, Melbourne, Australia: Linda Denehy and Thomas Rollinson.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Study tools in English and Vietnamese.
[PDF File (Adobe PDF File), 240 KB - humanfactors_v11i1e44619_app1.pdf]

References

1. Channa A, Popescu N, Skibinska J, Burget R. The rise of wearable devices during the COVID-19 pandemic: a systematic review. Sensors (Basel) 2021 Aug 28;21(17):5787 [FREE Full text] [doi: 10.3390/s21175787] [Medline: 34502679]
2. Atta RM. Cost-effective vital signs monitoring system for COVID-19 patients in smart hospital. Health Technol (Berl) 2022;12(1):239-253 [FREE Full text] [doi: 10.1007/s12553-021-00621-y] [Medline: 34786323]
3. Iqbal FM, Joshi M, Davies G, Khan S, Ashrafian H, Darzi A. Design of the pilot, proof of concept REMOTE-COVID trial: remote monitoring use in suspected cases of COVID-19 (SARS-CoV-2). Pilot Feasibility Stud 2021 Mar 05;7(1):62 [FREE Full text] [doi: 10.1186/s40814-021-00804-4] [Medline: 33673868]
4. O'Carroll O, MacCann R, O'Reilly A, Dunican EM, Feeney ER, Ryan S, et al. Remote monitoring of oxygen saturation in individuals with COVID-19 pneumonia. Eur Respir J 2020 Aug;56(2):2001492 [FREE Full text] [doi: 10.1183/13993003.01492-2020] [Medline: 32616588]
5. Seshadri DR, Davies EV, Harlow ER, Hsu JJ, Knighton SC, Walker TA, et al. Wearable sensors for COVID-19: a call to action to harness our digital infrastructure for remote patient monitoring and virtual assessments. Front Digit Health 2020;2:8 [FREE Full text] [doi: 10.3389/fdgth.2020.00008] [Medline: 34713021]
6. Local production and technology transfer to increase access to medical devices addressing the barriers and challenges in low-and middle-income countries. World Health Organization. URL: https://www.who.int/publications/i/item/9789241504546 [accessed 2023-12-07]
7. Mitgang EA, Blaya JA, Chopra M. Digital health in response to COVID-19 in low- and middle-income countries: opportunities and challenges. Glob Policy 2021 Jul;12(Suppl 6):107-109 [FREE Full text] [doi: 10.1111/1758-5899.12880] [Medline: 34230840]
8. Un K, Wong C, Lau Y, Lee JC, Tam FC, Lai W, et al. Observational study on wearable biosensors and machine learning-based remote monitoring of COVID-19 patients. Sci Rep 2021 Feb 23;11(1):4388 [FREE Full text] [doi: 10.1038/s41598-021-82771-7] [Medline: 33623096]
9. Rosic T, Petrina N, Baysari M, Ritchie A, Poon SK. Patient and clinician use characteristics and perceptions of pulse oximeter use: a scoping review. Int J Med Inform 2022 Mar 18;162:104735 [FREE Full text] [doi: 10.1016/j.ijmedinf.2022.104735] [Medline: 35325661]
10. Binyamin SS, Hoque MR. Understanding the drivers of wearable health monitoring technology: an extension of the unified theory of acceptance and use of technology. Sustainability 2020 Nov 18;12(22):9605. [doi: 10.3390/su12229605]
11. Morcillo LF, Morer-Camo P, Rodriguez Ferradas MI, Cazón Martín A. The wearable co-design domino: a user-centered methodology to co-design and co-evaluate wearables. Sensors (Basel) 2020 May 21;20(10):2934 [FREE Full text] [doi: 10.3390/s20102934] [Medline: 32455756]
12. Labrique AB, Wadhwani C, Williams KA, Lamptey P, Hesp C, Luk R, et al. Best practices in scaling digital health in low and middle income countries. Global Health 2018 Nov 03;14(1):103 [FREE Full text] [doi: 10.1186/s12992-018-0424-z] [Medline: 30390686]
13. Thai PQ, Rabaa M, Luong DH, Tan DQ, Quang TD, Quach HL, et al. The rirst 100 days of severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) control in Vietnam. Clin Infect Dis 2021 May 04;72(9):e334-e342 [FREE Full text] [doi: 10.1093/cid/ciaa1130] [Medline: 32738143]
14. Revans R. Action learning: tts originsnature. In: Pedler M, editor. Action Learning in Practice 4th edition. London: Routledge; 2011.
15. Brook C. What does action learning look like today? Action Learn Res Pract 2021 Jul 31;18(2):89-90. [doi: 10.1080/14767333.2021.1935058]

XSL•FO
RenderX

16. Boak G. Action learning and healthcare 2011–2022. Action Learn Res Pract 2022 Oct 17;19(3):251-268. [doi: 10.1080/14767333.2022.2133376]
17. Davis FD. Perceived usefulness, perceived ease of use, and user acceptance of information technology. MIS Quarterly 1989 Sep;13(3):319. [doi: 10.2307/249008]
18. Fishbein M, Ajzen I. Belief, Attitude, Intention, and Behavior: An Introduction to Theory and Research. Reading: Addison-Wesley; 1975.
19. Hong X, Zhang M, Liu Q. Preschool teachers' technology acceptance during the COVID-19: an adapted technology acceptance model. Front Psychol 2021;12:691492 [FREE Full text] [doi: 10.3389/fpsyg.2021.691492] [Medline: 34163416]
20. Lai P. The literature review of technology adoption models and theories for the novelty technology. JISTEM 2017 Jun 08;14(1):21-38. [doi: 10.4301/s1807-17752017000100002]
21. Tsai T, Lin W, Chang Y, Chang P, Lee M. Technology anxiety and resistance to change behavioral study of a wearable cardiac warming system using an extended TAM for older adults. PLoS One 2020;15(1):e0227270 [FREE Full text] [doi: 10.1371/journal.pone.0227270] [Medline: 31929560]
22. Legris P, Ingham J, Collerette P. Why do people use information technology? A critical review of the technology acceptance model. Inf Manag 2003 Jan;40(3):191-204. [doi: 10.1016/s0378-7206(01)00143-4]
23. Benbasat I, Barki H. Quo vadis, TAM? J Assoc Inf Syst 2007;8(4):211-218. [doi: 10.17705/1jais.00126]
24. Bagozzi R. The legacy of the technology acceptance model and a proposal for a paradigm shift. J Assoc Inf Syst 2007 Apr;8(4):244-254. [doi: 10.17705/1jais.00122]
25. Venkatesh V, Bala H. Technology acceptance model 3 and a research agenda on interventions. Decis Sci 2008 May 09;39(2):273-315. [doi: 10.1111/j.1540-5915.2008.00192.x]
26. Cooper A, Reimann R, Cronin D, Noessel C. About Face: The Essentials of Interaction Design. 4th edition. Indianapolis: Wiley; 2014.
27. Fan EMPJ, Ang SY, Phua GC, Chen Ee L, Wong KC, Tan FCP, et al. Factors to consider in the use of vital signs wearables to minimize contact with stable COVID-19 patients: experience of its implementation during the pandemic. Front Digit Health 2021;3:639827 [FREE Full text] [doi: 10.3389/fdgth.2021.639827] [Medline: 34713111]
28. Faria I, Gaspar C, Zamith M, Matias I, das Neves RC, Rodrigues F, et al. TELEMOLD project: oximetry and exercise telemonitoring to improve long-term oxygen therapy. Telemed J E Health 2014 Jul;20(7):626-632. [doi: 10.1089/tmj.2013.0248] [Medline: 24796364]

Abbreviations

HCW: health care worker
HTD: Hospital for Tropical Diseases
ICU: intensive care unit
LMIC: low- and middle-income country
OUCRU: Oxford University Clinical Research Unit
TAM: technology acceptance model
VITAL: Vietnam ICU Translation Application Laboratory

Edited by A Kushniruk; submitted 28.11.22; peer-reviewed by K Liu, M Harris, J Claggett; comments to author 15.04.23; revised version received 15.06.23; accepted 20.11.23; published 05.01.24.

Please cite as:
Luu AP, Nguyen TT, Cao VTC, Ha THD, Chung LTT, Truong TN, Nguyen Le Nhu T, Dao KB, Nguyen HV, Khanh PNQ, Le KTT, Tran LHB, Nhat PTH, Tran DM, Lam YM, Thwaites CL, Mcknight J, Vinh Chau NV, Van Nuil JI, Vietnam ICU Translational Applications Laboratory (VITAL)
Acceptance and User Experiences of a Wearable Device for the Management of Hospitalized Patients in COVID-19–Designated Wards in Ho Chi Minh City, Vietnam: Action Learning Project
JMIR Hum Factors 2024;11:e44619
URL: https://humanfactors.jmir.org/2024/1/e44619
doi:10.2196/44619
PMID:38180799

©An Phuoc Luu, Truong Thanh Nguyen, Van Thi Cam Cao, Trinh Hoang Diem Ha, Lien Thi Thu Chung, Trung Ngoc Truong, Tung Nguyen Le Nhu, Khoa Bach Dao, Hao Van Nguyen, Phan Nguyen Quoc Khanh, Khanh Thuy Thuy Le, Luu Hoai Bao Tran, Phung Tran Huy Nhat, Duc Minh Tran, Yen Minh Lam, Catherine Louise Thwaites, Jacob Mcknight, Nguyen Van Vinh Chau, Jennifer Ilo Van Nuil, Vietnam ICU Translational Applications Laboratory (VITAL). Originally published in JMIR Human Factors (https://humanfactors.jmir.org), 05.01.2024. This is an open-access article distributed under the terms of the Creative

XSL•FO
RenderX

Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on https://humanfactors.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Exploring a Gaming-Based Intervention for Unemployed Young Adults: Thematic Analysis

Ingjerd Jevnaker Straand[1], MSc; Asbjørn Følstad[2], PhD; Jone Ravndal Bjørnestad[1], PhD

[1]Department of Social Work, University of Stavanger, Stavanger, Norway

[2]Department of Sustainable Communication Technologies, SINTEF Digital, Oslo, Norway

Corresponding Author:
Ingjerd Jevnaker Straand, MSc
Department of Social Work
University of Stavanger
Kjell Arholms hus
Stavanger, 4021
Norway
Phone: 47 93 22 22 89
Email: ingjerd.j.straand@uis.no

Abstract

Background: Promoting positive psychologies that promote resilience such as a growth mindset could be beneficial for young, unemployed adults, as many lack the self-esteem and self-efficacy to cope with job search adversity. These young people may be reached at scale through the web-based delivery of self-administered positive psychology interventions. However, past studies report unsatisfying user experiences and a lack of user engagement. A gaming-based experience could be an approach to overcoming these challenges.

Objective: Our research objective was to explore how young, unemployed adults experience a positive psychology intervention designed as a game to extract learning and principles for future intervention research and development.

Methods: To respond to the research question, a team of researchers at the University of Stavanger worked with designers and developers to conceptualize and build a gaming-based intervention. Feedback from the users was collected through formative usability testing with 18 young adults in the target group. Retrospectively, recordings and notes were transcribed and subjected to thematic analysis to extract learnings for the purposes of this paper.

Results: A total of 3 themes were identified that pinpoint what we consider to be key priorities for future gaming interventions for unemployed young adults: adaptation to user preferences (eg, need for responding to user preferences), empathic player interaction (eg, need for responsiveness to user inputs and a diverse set of interaction modes), and sensemaking of experience and context (eg, need for explicit presentation of game objectives and need for management of user expectations related to genre).

Conclusions: Feedback from end users in usability-testing sessions was vital to understanding user preferences and needs, as well as to inform ongoing intervention design and development. Our study also shows that game design could make interventions more entertaining and engaging but may distort the intervention if the game narrative is not properly aligned with the intervention intent and objectives. By contrast, a lack of adaptation to user needs may cause a less motivating user experience. Thus, we propose a structured approach to promote alignment between user preferences and needs, intervention objectives, and gameplay.

(JMIR Hum Factors 2024;11:e44423) doi:10.2196/44423

KEYWORDS

positive psychology intervention; digital mental health; serious gaming; intervention design; research through design; gaming-based intervention

Introduction

Positive Psychology for Unemployed Young Adults

Young people who are not in education, employment, or training (NEET) comprise, on average, 12.8% aged between 15 and 29 years in the Organisation for Economic Cooperation and Development countries [1]. Studies show that negative self-perceptions and a lack of perseverance are barriers to successful labor market inclusion [2-4], as the new labor market requires highly skilled workers who are not afraid of change,

XSL•FO
RenderX

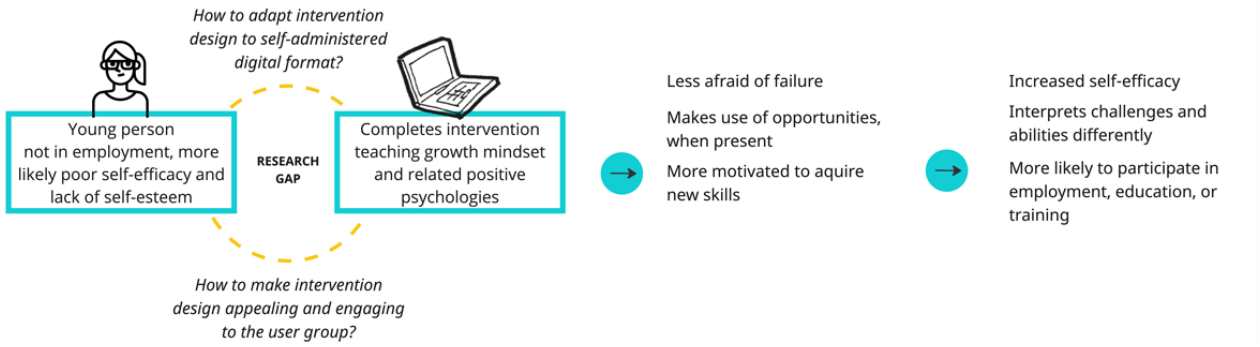
challenges, and acquiring new skills [5,6]. For young people with weak beliefs in their capacity to learn, this could be a major risk factor for labor market exclusion, and this may in turn impact their overall well-being. Several researchers have studied the relationship between unemployment and mental health. McGee and Thompson [7] found a relationship between unemployment and depression in young adults and suggested the use of psychological interventions for the young and unemployed. The Norwegian NEET group is more likely to be recipients of health-related benefits, have poorer mental health, and lower levels of education compared with the average of the Organisation for Economic Cooperation and Development [8,9]. A qualitative inquiry into young people's own experience of unemployment in Norway points to poor self-efficacy and lack of self-esteem that are reinforced through challenges and setbacks, even when these initially occur beyond the individual's control [10,11], such as when there are insufficient training placements on offer for the vocational school pupils, a problem leading to a relatively large number of unqualified school dropouts in Norway [9]. Thus, there is a substantial rationale for exploring further how the public can offer training, not only in job-seeking skills, such as curriculum vitae (CV) writing and gaining work skills, but also in building psychological well-being and resilience to cope with such setbacks and challenges [12,13].

In the context of a broader research project, the Career Learning App, our study investigates the design and development of a web-based intervention using positive psychology to achieve beneficial changes [14]. Our broader research idea is that young people in the NEET group, henceforth referred to as *young, unemployed adults* (for the sake of simplicity and to reduce stigma), could benefit from building confidence in the possibility of learning and improving. The research idea stems from a body of work that has demonstrated positive results from offering high school students self-administered positive psychology interventions (PPIs) centered on growth mindset and challenge-seeking behaviors [15,16]. A growth mindset is the belief that human capacities are not fixed but can be developed and increased in response to one's own efforts, good strategies, and help from others [17]. If a simple web-based PPI can influence high schoolers' mindsets in ways that lead to positive academic outcomes [15,16,18-20], then it could also likely be beneficial to the young unemployed, leading to changes in how they engage with their contexts. Despite this strong rationale for the applicability of PPIs to facilitate well-being and personal growth in vulnerable populations, they have only, to a limited extent, been tested and used in the context of unemployment [13,21]. However, we cannot simply apply the PPIs designed for educational contexts; they need substantial adaptation to be relevant or usable for this new target group of young, unemployed adults [17]. For instance, the school-related examples used within the PPI to make them relatable are not relevant to this new target population. Furthermore, there is a lack of shared context to piggyback on to deliver the intervention and ensure that users will adhere to it. Thus, there is a need to design and develop a web-based PPI designed specifically for young, unemployed adults and their context. If successful in user studies, a resulting intervention may be used in forthcoming large-scale randomized controlled trials in Norway.

Problems With Self-Administered Interventions

Self-administered web-based interventions have the potential to support well-being and positive health changes in a large number of people at a moderate cost [22]. However, this introduces new challenges, illustrated in Figure 1. First, there is the challenge of adapting current PPIs to self-administered digital formats that are fit for the purpose of the intended user population. Second, there is the challenge of user motivation to obtain the users to complete and adhere to the intervention [23,24]. Past research suggests that users are not interested in or do not enjoy using digital mental health interventions [25], suggesting a need to work on the actual interventions themselves to increase engagement and user motivation.

Figure 1. How mindset change may positively impact young, unemployed adults and the challenges of adherence, which we see as related to a research gap with a lack of knowledge of intervention designing.



Exploring PPI as Gameplay to Be Relevant for Young People

Past research suggests a need to adapt to the media preferences of young people and make the apps more visual and interactive to increase engagement and motivation among young people [26]. One possible approach to increasing engagement is to explore games and game elements. Starting from "where the young people are at" makes pedagogical sense [27], thus the application of game design is founded on young people's own interests as a way to foster engagement and learning of positive psychologies. Although play and games are not unique to young humans [28,29], the average age of video game players is now 33 [30]. However, playing video games continues to be popular among young people [30,31]. Interactive digital games are increasingly used for purposes beyond entertainment, as

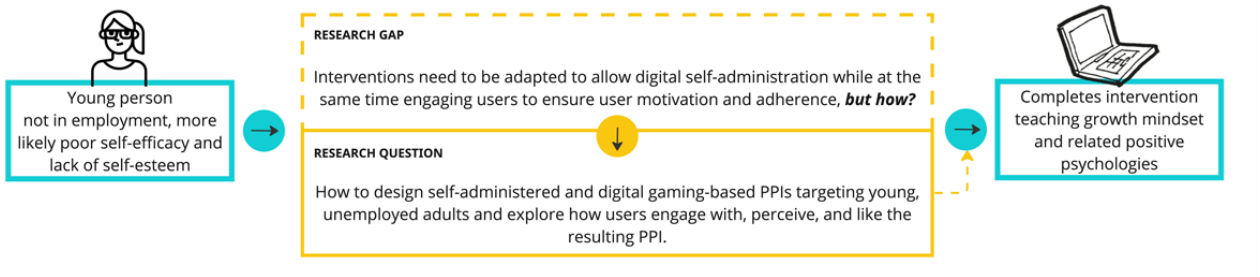
XSL•FO
RenderX

exemplified by the rise of health gaming apps for video gaming consoles. Game design elements are also increasingly applied to nongame contexts, for instance, by adding points and badges to nongame experiences, such as social media networks [32,33] or learning contexts [34]. When game design and game concepts are being applied for purposes beyond fun, they may be termed "serious games," "learning games," or "gamification" [35-40]. Game design applied to learning may be seen as a form of experiential learning (eg, learning-by-doing) [41,42]. Game design has been successfully applied to mental health interventions [43,44] and educational contexts [34,38] in the past. Game design offers an approach to creating engaging experiences. Engagement is a complex and ambiguous term [45]. Our use of the term is in the sense of "emotional involvement" as in offering a pleasurable experience [32] and to describe how motivational, usable, and acceptable [46] the game would be in the eyes of the target audience, because this could be an important predictor of adherence. In general, there is insufficient research on the application of gaming and gamification to mental health, particularly in the well-being domain [47], and we have not found empirical studies that pursue to gamify positive psychology targeted specifically toward unemployed young people. There were no available gaming-based PPIs that could be used for the purposes of this study.

Research Objective

Our limited knowledge of how to adapt the intervention from an educational setting to a game-based format suggested a highly explorative approach, where we identified the need to design a game to explore this topic and to overcome the gaps in knowledge summarized in Figure 2. Thus, the research objective and question of this study were how to design a self-administered and digital PPI in a gaming format targeting young, unemployed adults and to explore how they engage with the game and whether they like using it.

Figure 2. How the research question is linked with the research gap identified. PPI: positive psychology intervention.

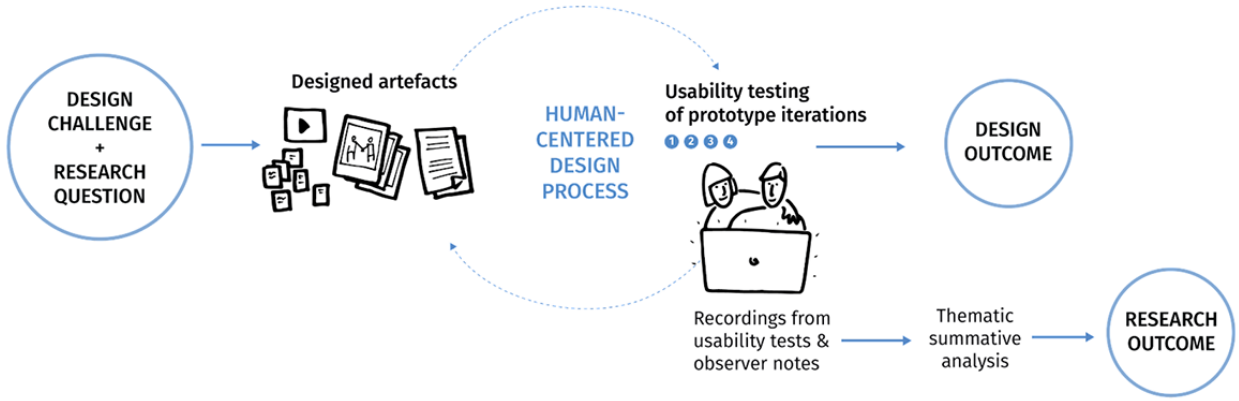


Methods

Setting: Learning in "Action"

To answer our research question, this study used a human-centered design process [48,49], an approach that allows input from target users during design and development. The human-centered design process is a form of research with design as the primary outcome [50]. The collection of feedback from users during design and development impacts not only the design of the game but also the production of knowledge. As such, it is a form of participatory action-research [51,52] where "doing research" and "doing action" happen simultaneously. In the design research literature, this may also be referred to as "Research through Design" [53,54], where knowledge is produced through the design of the artifact and through the experience of the artifact. At the end of the project, we analyzed user feedback data thematically [55] to extract the learnings and design principles for the potential application toward creating user-friendly and user-relevant gaming-based PPIs for vulnerable populations. Figure 3 summarizes the "Research through Design" approach of this study.

Figure 3. A Research through Design process with design outcomes and research outcomes.



XSL•FO
RenderX

Data Collection Through User Testing

We used testing with users to capture the user experience. Testing with users is important as a tool for obtaining the "design right" in game design [56]. Our project-specific usability study [57] was formative, and we were qualitatively evaluating early prototypes to answer how and why questions as a means to improve the design [58,59]. We aimed to capture participants' *thinking*, for example, opinions, reasoning, and attitudes toward the prototype experience. The various prototypes thus functioned as exploratory "hypothesis testing" [60,61] and as boundary objects [62-64] that framed conversations with end users. These sessions with the end users lasted from 45 to 90 minutes, where users were asked to briefly describe their background and interests, followed by open "think-aloud" questions [65] related to a prototype experience, such as "What do you think about what you see here?" "What do you expect will happen now?" and "What do you think this is?" Thus, the participants were encouraged to verbalize their thoughts and experiences. Afterward, the participants were asked follow-up questions that were equally open, such as "What are you thinking now that you have seen this?" and 'How would you describe this to a friend?" The objectives of usability testing were to collect feedback related to broad aspects of the intervention experience, namely its (1) engagement, (2) relatability, (3) understandability, and (4) potential for improvement.

Sample and Recruitment

The NEET group includes anyone not in economic activity from the age of 15 to 29 years [1]. The European Foundation for the Improvement of Living and Working Conditions has defined 7 subgroups of NEETs [66], ranging from the "classically unemployed" to young people who are caretakers, unable to work, or simply listed as "inactive." How long an individual remains a NEET also varies significantly. The diversity of the population is not necessarily problematic for our study; when designing with users, one would usually strive for variance [49,67] rather than representativeness. Our study is exploratory and does not require large samples [68]. Our estimate required 20 users; however, this sample size was highly approximate, in line with qualitative studies in general [69,70]. Pragmatic needs in the design process guided the number of participants to a large degree and not, for instance, theoretical saturation.

We recruited from the Norwegian Labour and Welfare Administration (NAV), the national welfare institution of Norway that pays out unemployment wages and social support, and from a regional Individual Placement and Support (IPS) program, which offers placement support to young people with first-episode psychosis. We also recruited through the user testing platform, Teston (UserTesting). Our inclusion criteria were an age range of 18 to 29 years, with a "NEET background," and because of the language in the game prototype, living in Norway and speaking Norwegian. We say "NEET background" and not "NEET status" because our participants from the IPS program were no longer in the NEET group by definition. We made a deliberate choice not to exclude based on the length of NEET status and unemployment. Although those who are entering the NEET group in the short term, the "in-betweeners" [66], often find new employment without assistance [71] relatively quickly, even a short time out-of-work may increase the risk of exclusion [72]. We did not recruit participants on permanent disability allowance. Using multiple channels enabled quicker recruitment during the COVID-19 pandemic and increased the variance in our sample (as desired), as young people without the rights to receive unemployment benefits have fewer incentives to register with NAV [73]. We did not compare the experience of the intervention based on the recruitment channel because the groups were overlapping and experienced varying prototypes depending on the stages in the design process.

Participants

In total, 18 participants (12/18, 67% females, 6/18, 33% males) took part in the study during the 21 testing sessions; thus, some participants were involved more than once. Recruitment was particularly challenging because of the COVID-19 pandemic, and we found that it was difficult to recruit young, unemployed men. Remote participation through web-based technologies, such as the Zoom (Zoom Video Communications) platform, enabled the study to continue during the lockdown. We also experienced that this user group continued to prefer remote participation, even when restrictions were lifted. Table 1 summarizes the participant statistics and format of the usability test.

XSL•FO
RenderX

Table 1. Age distribution of participants and format of user testing.

Recruitment channel	Format user testing	Age (years)[a]
IPS[b]	In person+Remote	19
IPS	In person+Remote	28
IPS	In person+Remote	18
IPS	In person+Remote	27
IPS	In person	22
IPS	Remote	19
IPS	Remote	18
IPS	Remote	23
Teston	Remote	18-29
Teston	Remote	18-29
Teston	Remote	18-23
NAV[c]	Remote	21
NAV	Remote	18
NAV	In person	21
NAV	Remote	21
NAV	Remote	18
NAV	Remote	18
NAV	In person	22

[a]For 3 (17%) of the 18 participants, we only had an age interval provided to us.

[b]IPS: Individual Placement and Support.

[c]NAV: Norwegian Labour and Welfare Administration.

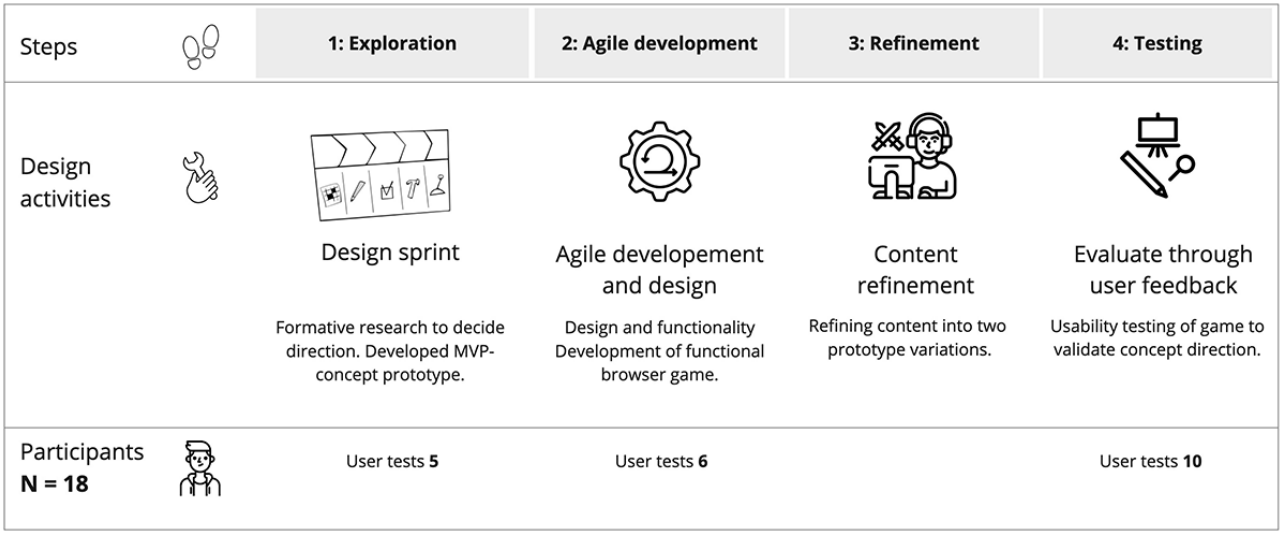
The Design and Development Process

Researchers at the University of Stavanger worked with designers and developers from a consulting company and potential future end users in an agile [60] and human-centered design [48,49] process. The objective of the process was to create an enjoyable gaming-based PPI targeting unemployed young adults. This process took place in 4 steps over approximately 9 months, from 2019 to 2020. In the first step, (1) *design exploration* in the form of a design sprint, a 5-day design and prototyping process [61], produced a minimum viable product of a gaming concept that was tested with 5 participants. Following a brief period for planning and procurement, we moved on to (2) *agile development*, consisting of 3 sprints, each lasting about a month. During this step, usability tests were conducted on 6 participants. This was followed by (3) *refinement* of content and prototypes, with involvement from behavioral intervention researchers to further develop and "add-in" the necessary intervention content. Finally, the prototypes were evaluated using (4) *testing*. In this step, 10 participants participated in usability testing and provided feedback on the final set of prototypes. Figure 4 summarizes the stepwise design and research process.

The intervention content needed to be adapted to be meaningful and relatable to the user group [74]. The basis for our gaming-based PPIs were growth mindset interventions from the "National Study Learning Mindset" [16] and its translated Norwegian version, "U-SAY" [5,15]. These are interventions that target high school students [15,16]. We added selected parts of cognitive behavioral therapy [75], specifically management of negative emotions, panic, and anxiety, to offer a more productive interpretation of stressors [76] that may occur during job search adversity [77,78]. During the first few days of the initial design sprint, a gaming concept, VitaNova, was developed where players can build a "new life" in a fictional narrative.

XSL•FO
RenderX

Figure 4. The stepwise design and development process.



The Gaming-Based Intervention

The VitaNova gaming concept is a one-player fantasy game where the user plays a "no name," an android character that can take on any skin to morph into another character with their skills and abilities. In particular, the player can choose between being Noomi and Twizzlesprock. However, as we learn through the game, your character has a backstory as the male character Abel, a former engineer and an outcast who sees himself as a failure. The game is designed in 3D and split into episodes (missions). The game starts with very little information and instructions, landing the user right into action. This was a design tactic to spark curiosity and make the users intrigued by the game so that they would want to explore it further. As a player, your first task is to escape from captivity, and then, gradually, more and more information is being revealed to you as the game progresses. Refer to Textbox 1 for an overview of the game narrative.

The visual design of games can influence how motivational and acceptable they are to the target population [46]. Therefore, to make the game look polished, cool, and professional and to keep users immersed and engaged [79], emphasis was placed on 3D design and detailing. The game was divided into episodes, or missions. The first 3 episodes were developed into a nearly fully functional game in Unity WebGL, a platform for building 3D games that can be used in a web browser. During user testing, we also showed the prototypes and the wireframes that were made in Figma. Figure 5 shows the prototype iterations of the game design.

Psychological content and tasks are entered into the gameplay to foster psychological well-being, teach a growth mindset, and offer psychoeducation and mental health tips. Some of this is interwoven into action in the form of interactive quizzes, dialogues, or other forms of interaction, such as a CV builder applied within the game. This was intended to be transferable to the end user situation to increase relevance, although the acquisition of such practical skills was not a target of the intervention. Furthermore, there was also psychological content that was external to the gameplay, such as embedded videos. When using externally sourced content, this was implemented in the game as "ruins from the past," which the player could "find" in the game. The player would need to watch this content and use the information provided to complete the challenges and the in-game quizzes. Upon completing an episode of the game, the player was requested to write an answer to a reflective question where the user should answer as himself or herself, to encourage internalization of the messages that had been taught in the intervention through self-persuasion [80], and to transfer learning to the user's own situation. We have included a further description of the prototypes in Multimedia Appendix 1.

XSL•FO
RenderX

Textbox 1. Description of the game narrative. (The 2 final missions were not included in the user testing.)

1. Introduction

Wake up in the trunk of a moving vehicle. Use hacking skills to hack the lock. Find a tavern and interact with a bodyguard who refuses your entry as "no name." Find the Noomi skin and power unit and turn them on to enter the tavern.

2. Tavern

Interact with Griff and Mia at the tavern to learn about this world. Try to receive help. Tell them that you must go to an old public office to pick up energy bars if you are going to receive any help.

3. Learn

Find an abandoned public office where, among other things, you will discover many pieces of ancient psychological knowledge as ruins from the past.

4. The sidekick

Use newly acquired knowledge to help out the depressed and anxiety-ridden droid Griff, who now becomes your sidekick, and help Mia the bartender.

5. The bully

A bounty hunter is out looking for you to receive a reward from the boss at BetterJu Janus. Threatens your new friends at the tavern. Why are they after you?

6. The chance

You discover that the only way to obtain all the answers is to try to join into BetterJu. Your friends tell you of a job posting that is open. Interactive curriculum vitae and job application process.

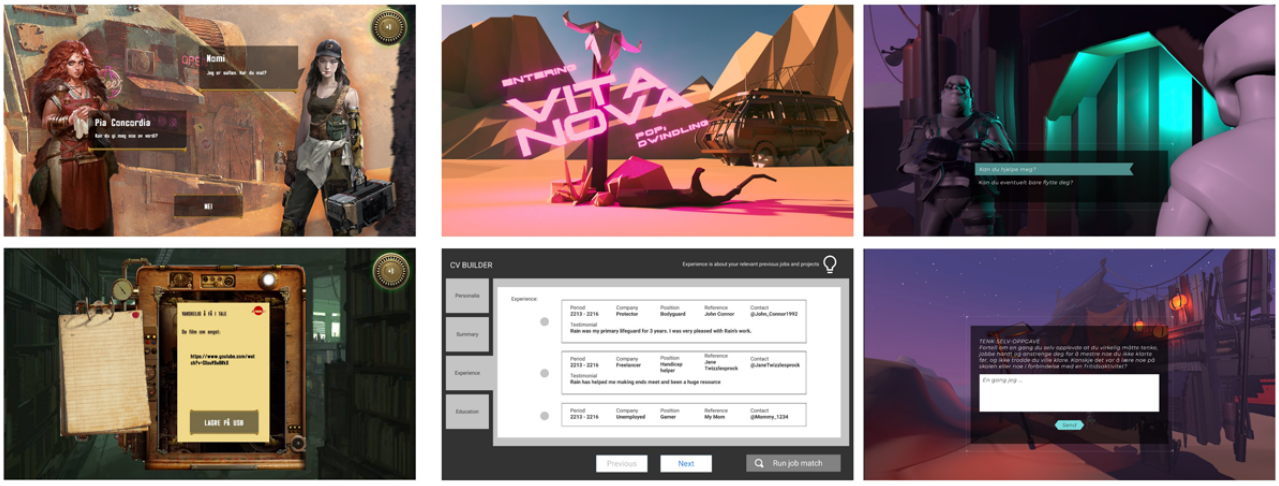
7. The job interview

The job requires a different set of skills, which you acquire through entering the skin of Twizzlesprock. Job interview at the company BetterYou as Twizzlesprock.

8. The escape

Discover who you really are from the overhearing conversation between your new boss, Janus, and a droid. Find your ex-girlfriend, who has been trapped but confirms your true identity. You both escape from the evil boss. The end.

Figure 5. Prototype iterations of VitaNova.



Analysis

The purpose of the usability testing was to synthesize findings that led to improvements and changes in the designed outcome. Retrospectively, we also conducted thematic analysis with the steps from Braun and Clark [55,81] as a practical guide: (1) data familiarization, (2) initial code generation, (3) search for themes, (4) review themes, (5) define and name themes, and (6) produce reports [81]. All 21 usability-testing sessions were recorded. The 7 most comprehensive usability tests were transcribed verbatim. The remaining data were analyzed based on recordings, researcher notes, and memos. Specifically, we used the transcribed data as our starting point and went back to recordings and memos to review codes and themes. The analytic process was iterative and creative, where we often moved back and forth between the data and codes [82]. All authors independently familiarized themselves with the data. Author 1 started with coding using the qualitative analysis software ATLAS.ti (ATLAS.ti Scientific Software). As a group, we discussed the findings and the initial codes in a workshop before moving over to paper-based coding and printing quotes from participants organized on large paper sheets. We used

XSL•FO
RenderX

diagramming, both digitally in Miro and in pen-and-paper sketches, to iterate themes and review codes in consensus meetings. Findings and concepts were discussed with other researchers, some of whom had acted as observers for the user testing or had watched recorded sessions. The quality of the analysis was ensured by researcher reflexivity, end user involvement, and method triangulation. *Researcher reflexivity* concerns activities that consider how researchers might have informed the research or biased outcomes [69]. Reflexivity was enabled through critical discussion of assumptions, themes, and codes in the team of researchers. The team also involved researchers not involved in the user testing or in the design process as an approach to validate the analysis based on these methods. *End user involvement* was supported through the iterative design process, where participant perspectives were sought at different levels of concept and design maturity. Specifically, we found it valuable to involve users from different recruitment sources (IPS, Teston, and NAV) to strengthen the credibility and transferability of the findings. *Method triangulation* was conducted by applying different approaches to design and user involvement at different phases of the process, allowing the assessment of themes or constructs from different perspectives. In particular, including data from the different phases of exploration, design, and evaluation was found to strengthen the credibility of the findings.

Ethical Considerations

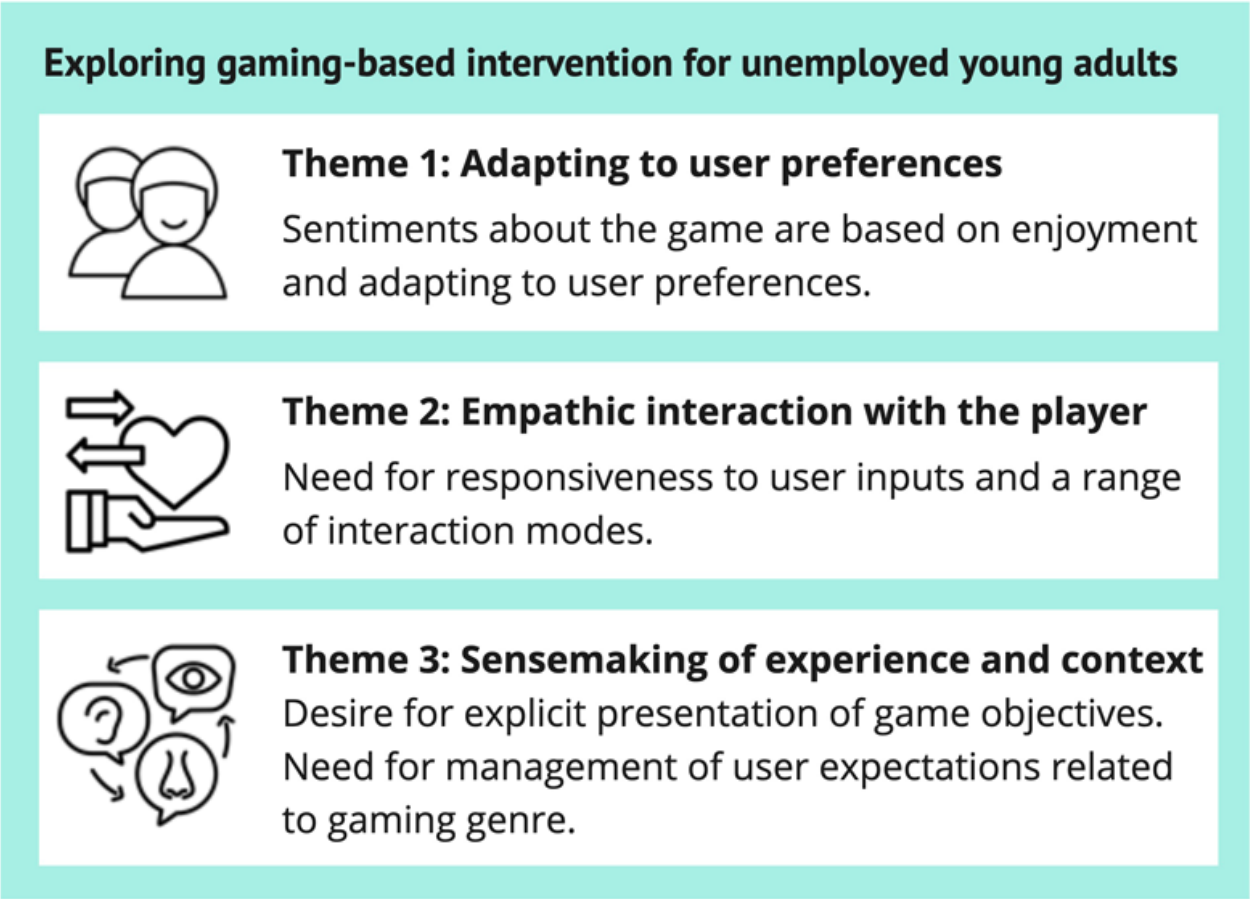
All participants provided explicit and written consent to participate in the study and were rewarded with gift cards (approximately US $30/session) for their participation. The study was evaluated and approved by the Norwegian Centre for Research Data (approval number 131074) and the regional committees for medical and health research ethics in Norway (approval number 42128).

Results

Overview

As a tool to motivate and help young adults engage in work or education, the idea of an interactive and digital game was regarded by all participants as a "good idea"; it was described as "cool," "unexpected," and "motivating" upon first impression. Upon closer experience with the different gaming prototypes, we received different and more specific feedback. A total of 3 themes were constructed by the researchers through active analytic engagement with the data [55]. The themes pinpoint what we consider key priorities for future gaming interventions for unemployed young adults: (1) adapting to user preferences, (2) empathic player interaction, and (3) sensemaking of experience and context. Refer to Figure 6 for an overview of the themes and their key characteristics.

Figure 6. Themes and theme characteristics.



XSL•FO
RenderX

Theme 1: Adapting to User Preferences

Sentiments about the game are based on enjoyment of the game experience, the genre, and whether the game meets user preferences, either by being targeted toward them or by allowing for experience customization. The participants expressed positive feelings and excitement related to the game. It was described as much more of an *actual game* than was expected:

> *I liked it; it was unexpected.* [P3]
>
> *Seems like a cool concept. Never heard of it before.* [P17]

The participants pointed to the use of humor and compared the game to pure entertainment media, such as commercially available games, or other parts of popular culture, such as films or television series:

> *It was actually pretty exciting [laughs]. Kind of funny, considering that you have included the welfare administration here.* [P6]
>
> *The story was funny, it seemed a bit like a video game.* [P10]

The game concept was further described as "something to do" or "something to allow the time to fly," as pure entertainment or for relaxation purposes. The relaxing features were described as something that could make one more receptive:

> *If it's like that you get new assignments once a week, then it can seem exciting. At least that is something to do.* [P7]
>
> *So you are pretty relaxed when you get these questions, so then it is probably a little easier to answer... a little easier to reflect over this.* [P4]

Others felt that the gameplay story could motivate them to do something by creating a sense of urgency, where the story would drive them to make more effort. A participant wanted proof within the game that it would be worth the effort to perform mundane tasks, such as updating their CV:

> *That's how it is in life too: You have to make an effort yourself to get ahead; when you have to do something here [...], then you have to hurry, because someone is after you. Then you get the adrenaline to do it.* [P16]
>
> *If there's something in the game that can prove to me that it's worth it, like writing a resume is worth it....* [P18]

However, several participants expressed uncertainty about whether the game would be "something for them," suggesting either other preferences or not quite the right conceptual fit. In particular, the concept likely needed "aging up," as it was perceived as something for a younger population. Furthermore, we learned from the participants, who spent more time gaming, that they had started out playing "adventure games" and "role-playing games" when they were younger, but that they had since moved on to playing "first-person shooter" games or other kinds of games:

> *It's not a game I would have bought in the store.* [P15]
>
> *I think if I had been younger then, yes. Because I have a lot of different types of games that I like to play, and now I like to play games where you shoot people, but earlier I liked playing games like that, where you follow a story, for example, it's very different in... it's very different for people what kind of games they like to play.* [P2]

Designing for mobile phone use was considered important and was brought up as an improvement suggestion by nearly all users. Participants expressed that mobile phone use would make it easier to meet their own user preferences or the user preferences that they expected other young people to have:

> *I would have chosen an app or a mobile game. Or a course, if it was on [a] mobile [phone].* [P7]
>
> *Not everyone has a PC with them everywhere, so I wondered if this was on mobile.* [P17]
>
> *If the game could be turned into a mobile game, that would be better.* [P1]

Adapting to user preferences could also mean designing a customizable or more personalized experience. In particular, the choice of characters in the game is usually an arena for customization and personalization. One user commented that she would prefer to customize a character by selecting hair length, and body shape, etc rather than choosing between predefined female and male characters:

> *if you can choose male or female or... you don't have to have these two, but you can choose what that person looks like. Since now a lot of young people... there are some young people who don't want to be a man or a woman, so I think you... it's very smart to make something like 'do you want long hair' 'do you want short hair' 'do you want...' [laughs]* [P2]

Others mentioned different strategies for choosing characters by either choosing a character that resembled themselves or identifying with them. Others would deliberately choose the extreme opposite of themselves. As one participant mentioned, "if your choice was the big male character, then you were likely 'more vulnerable on the inside' and 'in need of protection.'"

Theme 2: Empathic Interaction With the Player

This theme describes the need for responsiveness to user inputs and the desire for a range of interaction modes. Although initially intrigued and enthused by the game design, users quickly became disappointed by the lack of functionality. Thus, this theme is based on the need for empathic interaction with the player in the game, where the game needs to take the user seriously by being responsive to user inputs, thereby allowing for actual contribution to the experience:

> *It's very... you can see very easily that your answers don't make much of a difference. It doesn't matter what you choose. And I think if you're going to have a game like this, you have to have a little change in what you say, how will it affect the game.* [P2]
>
> *No, it's just that I want to see that the people you're talking to have something else or something more to*

XSL•FO
RenderX

> *say. That they have other reactions than just "no [laughs]."* [P2]
>
> *It was, it's not a challenging game, it's not a difficult game... the game itself [laughs]. It's designed to be clicked through. You don't need to spend a week on it.* [P3]

Participants find the graphics visually appealing, although they do not feel that this is the most important aspect, saying that how the game works and how exciting and entertaining the game is are the most important parts:

> *I think it looks pretty nice. There are many different types of games that can be in like... many different types of 'art' [styles] and so, yes, there are many games that can look one way, but can still be really good, but some games that look really nice, can be boring. So, don't worry about how the game looks, but how the game works, that's very important.* [P2]
>
> *Visually it was very nice... The story I am more unsure of....* [P12]

Some of the feedback indicates that our game was perhaps not fully developed as a game with the necessary combination of rules, goals, feedback, fantasy, and fun [32]. Participants wish for a wider range of interaction modes, such as moving more freely around in the game, having more challenges and tasks in the game, and having different ways of interacting with the characters. Users expect game-like interactions, not just choosing answers, reading, and writing:

> *It is unusual for me that you cannot move around [in the game].* [P4]
>
> *But I think that it would be a bit boring if it was just like that you had to read, and then click to choose answer options. [...] if I was sitting at home, and this was something I had to go through every week[...] then I'd just click quickly through it. And then I hadn't properly read what it said.* [P7]

To support the learning objectives of the game, players had to answer reflective questions at the end of each gaming session, where they answered as themselves and not as the gaming character. This felt a bit "off" to the participants. Furthermore, several participants expressed a problem with articulating answers to those kinds of questions, expressing that they would not know what to write when asked:

> *I didn't expect that an assignment came up where you have to write about an experience from reality, sort of, which seems a little unusual to me since you are sitting inside this alternate world. But I had probably only written something about skateboarding. But it was very unusual for it to be like that.* [P4]
>
> *I don't know... at least I struggle a lot with tasks like that[...] I probably wouldn't have written anything here.* [P5]

Although not as exciting as hoped for, the challenges and tasks in the game can still provide the user with a sense of achievement and act as an awakening for new thoughts; if not for them, then perhaps for someone else:

> *You kind of get a little more confidence in yourself then. That you have actually managed something.* [P6]
>
> *...after all, it raises thoughts and yes... new ways of looking at things, I think. that it can start something in someone.* [P7]

Theme 3: Sensemaking of Experience and Context

This theme comprises a desire for an explicit presentation of the game objectives and a need to manage user expectations related to the gaming genre. The game was described as "cool but confusing:"

> *Uh, well, it seems kind of cool, but it was a little hard to understand, I felt.* [P5]

Many participants mentioned *sensemaking* or lack of understanding in some form or another; they struggled to understand the point, the objective, or the mission to be completed in the game. Some participants pointed to a lack of logic or strangeness in the storyline and over-the-top reactions to what they perceived as minor happenings:

> *So people want to buy parts of dead people so they can look how they want? Hum. That's a very strange concept! [laughs].* [P2]
>
> *She is stabbed now! ... And the taverna is burning. That was over-the-top. She just came for some food. This is over-the-top.* [P12]

The intervention messages in the game were not perceived by any of the participants. They were uncertain about what they were learning from the game experience. They were focusing on the details of the game narrative, trying to make sense of that, and, thus, the intervention part seemed "part of the fiction" and not clear what this was meant for:

> *Didn't learn anything. Well... I learned that there can be different ways to solve things, but I didn't really learn anything[...] It was a bit difficult to understand the whole story, that is the whole thing.* [P1]
>
> *Lots of talk about the brain, that the brain is a muscle, but don't know what it can help with, it doesn't make sense.* [P2]

The gameplay added complexity and was confusing to the participants. For most participants, there was a desire for an explicit presentation of game objectives, both in terms of what it should ultimately achieve for the end user (as an intervention) and what the objectives in the game narrative are. However, other participants felt that this uncertainty was part of the excitement:

> *...I should have known a little more what the goal was and what the meaning behind the game was. Because it seemed a bit like that, yes... a bit out of the blue. And you didn't quite know what an anonymous person was and whether this was the future or whether this was a completely different world.* [P7]
>
> *It was very interesting. It was very unusual for me with that kind of game. But I liked how it was. And you didn't have very much information about what*

XSL•FO
RenderX

> *you were doing so you kind of had to find out a bit about the skins and such. And I liked that.* [P4]

One user expressed explicit concern about how relatable the contents could be if you use a context that is far removed from everyday life:

> *If it becomes too sci-fi, I think it might be difficult to transfer to reality.* [P18]

Furthermore, we also identified a need to manage gaming genre expectations. The participants expressed a preference for certain genres over others; it may be difficult to cater to different preferences in terms of what games they like best to play. There are also certain expectations connected with different gaming genres that we were not so aware of in the research team, where users were trying to make sense of the game prototypes *in relation to* established genres, with expectations of gaming interaction to be similar to games in that genre. The participants asked us about the game in relation to genre concepts such as "open world," "adventure," and "role-playing" games:

> *So, I have a question, is this an open-world type of game where you go out to different places and pick up things or is this a text where you just follow what happens in the story?* [P2]

> *I have a question: Is this open world—or just to follow a track, like?* [P17]

It seems that the game genre was not clear to the participants, who pointed to different features of the gameplay prototypes that would take the game in different genre directions.

Discussion

Principal Findings

This study has used an iterative design process with active participation from potential users to develop an interactive game that aims to be user-friendly and engaging to be able to provide a vulnerable population with positive psychologies. As pointed out by past research, there is a strong rationale for promoting psychological well-being, for instance, to improve resilience [83] in the face of setbacks and challenges that occur as part of job search and being "out-of-work" [12,13,78], and may thus alleviate suffering [84]. A total of 3 themes were constructed from the user-based research that occurred through the design process of the gaming-based intervention: (1) adapting to user preferences, (2) empathic player interaction, and (3) sensemaking of experience and context. In the following section, we discuss the themes and how they could potentially be applied as designing principles for future self-administered gamified PPIs. Thus, the study sheds light on the application of game design for PPIs that aim to promote well-being and increase challenge-seeking in young, unemployed adults.

Comparisons With Previous Work

Adapting to User Preferences

This study expands the knowledge found in other studies on PPIs for young people, where the need to offer interactive, visual, and more engaging experiences has been identified [26,85,86]. Most participants expressed uncertainty about whether the game would be "something for them," suggesting other preferences or not quite the right conceptual fit with their preferences. We interpret this to mean that there is a need to consider gaming genres and user preferences specifically, where a more refined user segmentation may be necessary [87,88]. For instance, in our group, some participants said that they only had an interest in certain kinds of games. Future studies could consider a more fine-grained targeting strategy based on preferences and interests and not simply age and employment status. A possibility would be to segment the population based on player types [89,90] or motivation [89,91-93], and to think more carefully about user preferences for different gaming genres before choosing a concept. In our study, we found that the selected game genre was perceived as engaging "not to them," but "someone younger." This was particularly true for the active gamers, who found the genre to be immature. Furthermore, most participants stated that they would have preferred a game designed for mobile use, indicating another kind of context of use than our initially planned use on PCs at home.

Empathic Player Interaction

Inside the game experience itself, the participants in our study were disappointed by the lack of features, functionalities, and opportunities to influence what was going to happen in the game, for example, player autonomy. A lack of autonomy may cause a more negative interpretation of an experience [94]. We interpret this as underdelivery, partly because of the overpromise of the first impression and the esthetics of the graphics [95,96]. The participants expressed bleakly that in-game actions "do not matter" because, as players, they experienced an insufficient influence on the string of events in the game. Autonomy is an important motivator in self-determination theory (SDT) [97] and a lack thereof may contribute to reduced user motivation [13,35,97]. In part, the lack of autonomy and interactivity was caused by the requirements for a structured intervention set by the broader research project; each player needed to experience the same sequence of events. However, even within this frame, the game should be built to cater for somewhat more variation and focus on the interaction between the game and the player to meet user expectations. There is also a more specific need to be aware that within this target group of young, unemployed adults, many may feel *in general* that what they do does not matter [98]. Gameplay with insufficient df may unintentionally reinforce that message.

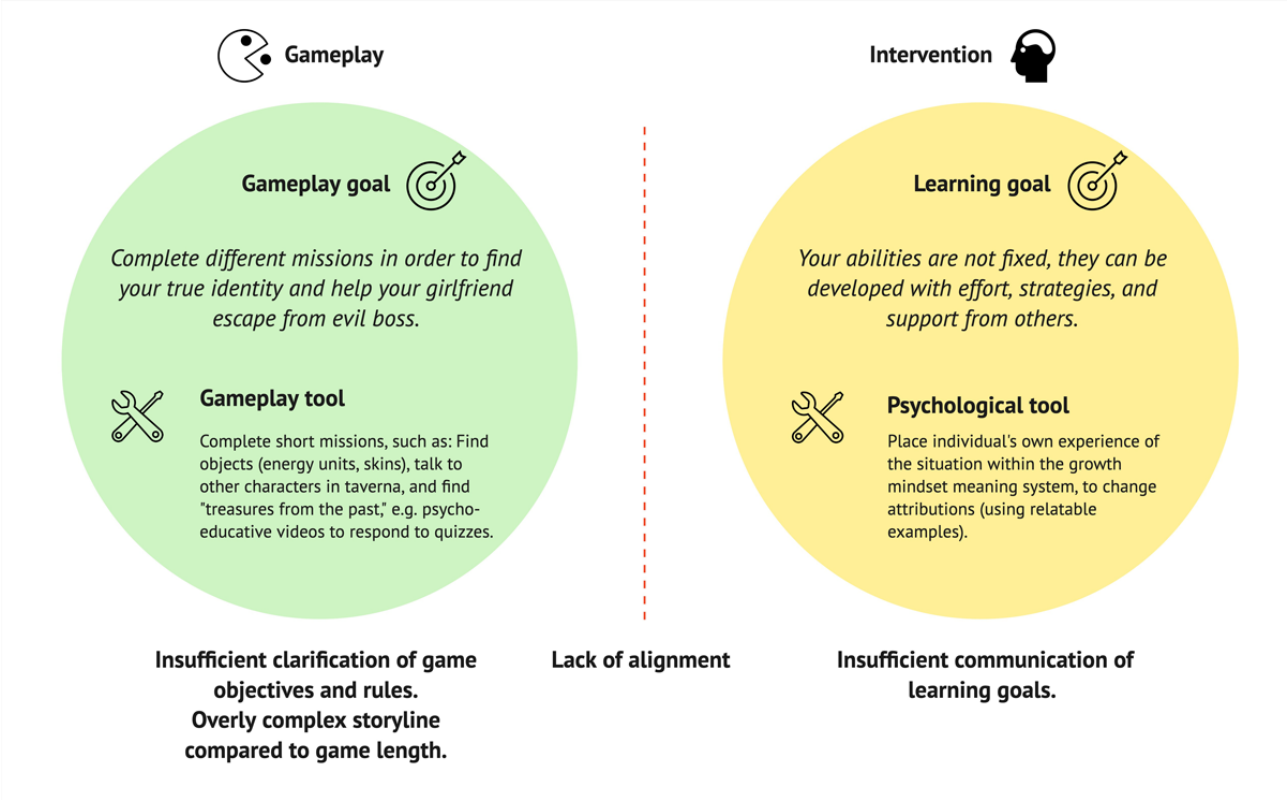
Sensemaking of Experience and Context

Participants had trouble making sense of the experience: our prototypes did not meet the participants' expectations for genre, not fitting with role-playing games based on narratives and dialogues or open-world type games, where you move around freely in a 3D world to pick up items and battle with other characters. Furthermore, the game objective and rules were either not clearly presented to the participants or did not cater to sufficient player-game dialogue and manipulation of the experience. We found that there were tensions between the gameplay and the messages of the intervention, which could undermine the intervention and potentially threaten its effectiveness. This finding is depicted in Figure 7. The PPI

XSL•FO
RenderX

gameplay had a complex storyline, which confused the participants and made them miss out on their learning objectives. Past studies have pointed to psychological affordances and the importance of a "fertile soil" to make positive psychological interventions more likely to work [99-101]. In instructional or serious games used in education, Young et al [38] concluded the need to ensure that game objectives and learning objectives correspond and, further, that an overly complex gameplay can lead to misunderstandings and interfere with understanding. This seems transferable to gaming-based interventions. Other authors have referred to this as "relevant narrative," which states that the narrative of the game should be relevant to the subject matter [102]. The choice of gameplay as a strategy for creating engagement for an intervention introduces a new context, which becomes the background for interpreting the messages of the intervention. The game design concept should be selected carefully and tested early with inexpensive methods, such as roleplay or paper sketches [32,56], to explore whether the gameplay is supportive of the intervention. In VitaNova, the gameplay goals implicitly reflected the learning goals, as the development of abilities was presented through the completion of in-game missions. However, because these learning goals were not explicitly communicated, the effectiveness of the intervention depended on the users themselves seeing the connection and transferring this knowledge to their own situation. Combined with the lack of clarity of game objectives and rules as well as an overly complex storyline, this led to confusion.

Figure 7. Mismatch of goals between intervention and gameplay, in combination with insufficient clarity overall.



Practical Implications for Future Gaming-Based Interventions

When revisiting the 3 themes and comparing them to related work, there appears to be a similarity between the identified findings of this study and the 3 basic psychological needs in SDT [103], which are: needs for autonomy, relatedness, and competence [92]; refer to Figure 8. As such, this study provides a form of bottom-up support for the usefulness of these constructs in designing and evaluating future gaming-based PPIs to understand how they might be more motivating to the user [92]. Further research should also investigate how and if a gaming-based PPI experience that does satisfy the relevant needs of autonomy, relatedness, and competence may contribute *in itself* to positive psychological outcomes for this population, as has been suggested [13].

Striking the right balance between learning and fun is a significant challenge, along with producing a relevant narrative [102] that supports intended learning. Ferrara [32] suggests a strategy for identifying the "gameness" that already exists in a context or situation rather than trying to *tack it on*. This may make it easier to transfer learning from the gaming space to everyday life [56,104]. However, moving from an idea to a game that works conceptually is challenging [105], and good intentions may be undermined by a seemingly fun yet unfit idea or concept, for example in the case of Disney and their first version of the game and exhibit "Habit Heroes," intended to support healthy eating but rather reinforced stereotypes and made children feel bad about themselves [106]. Choosing an approach that "gamifies life" should thus be done with empathy, care, and frequent testing with users to avoid banalizing the situation and experiences of a vulnerable population, such as the young and unemployed. As such, a human-centered design approach is ideal because it starts with empathy [48]. However,

XSL•FO
RenderX

frequent playtesting [56] and usability evaluation [57] are also needed to reduce the risk of developing a concept that is not engaging with the intended audience [88] or that undermines or does not foster learning. Established game genres and concepts could be used as inspiration in early explorative ideation. The characteristics of existing games may be viewed as opposing values on a spectrum [32], and by imagining what the game-based PPI would look like in the form of existing game genres, a large volume of different ideas can be formed that may be tested early for fit with the PPI objectives and user preferences.

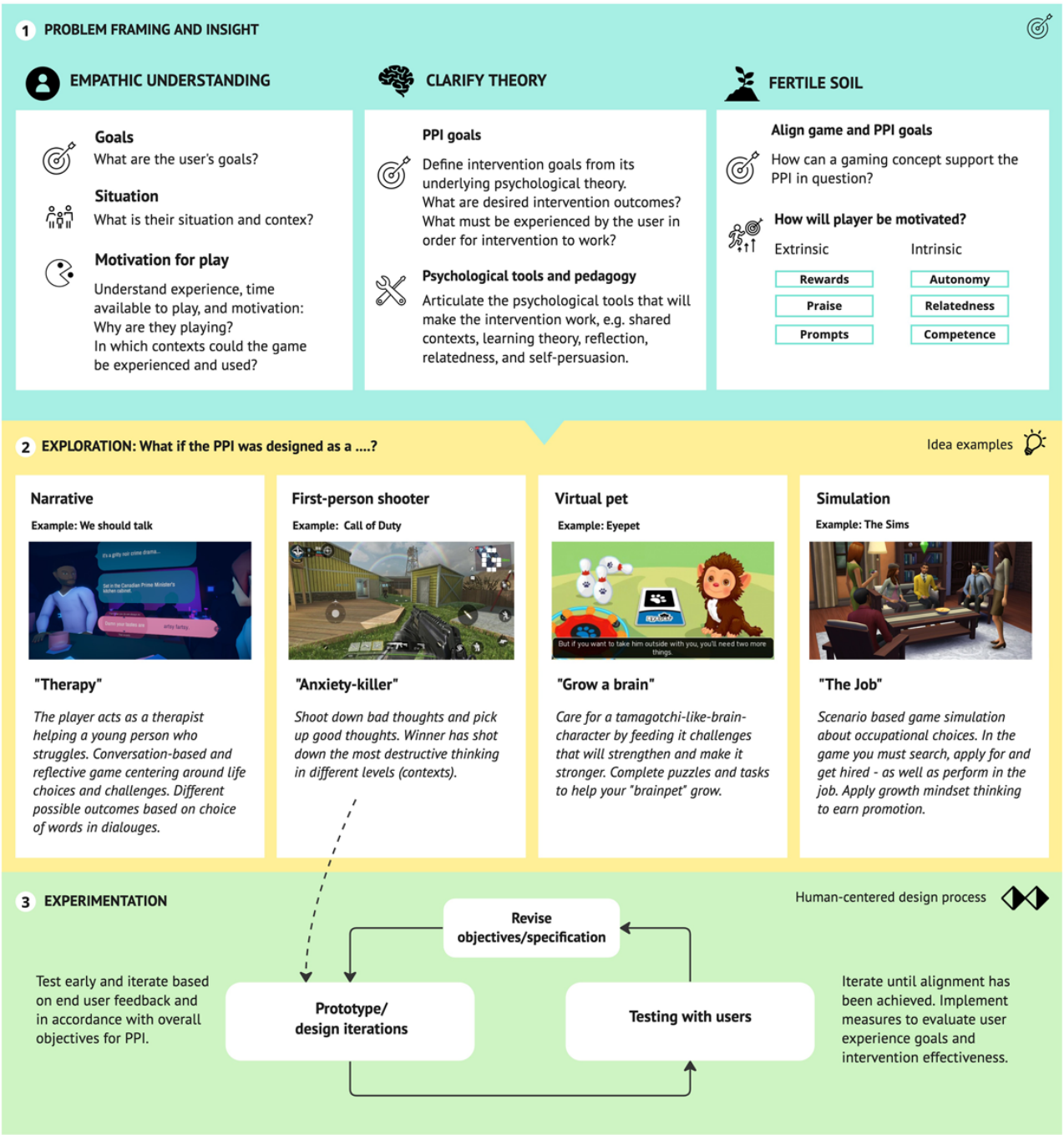
In Figure 9, we propose a broad but structured approach for how game-based PPI exploration may be executed, based on the lessons learned from our study and the discussion points in the preceding section. In this approach, insight into the user, context, gaming preferences, and gaming interests frames the design problem. It is also necessary to establish a clear and precise definition of PPI, including its underpinning mechanics, theories, and strategies that can help make the intervention effective. An alignment between the 2, forms the necessary "fertile soil" for the intervention game, where we ask how gameplay may support both user preferences and goals and PPI goals. Next, we propose working with existing genres and games to quickly generate many different ideas of what our gaming-based PPI may look and feel like. Promising concepts should be evaluated against relevant criteria, such as gameplay and user experience objectives and PPI objectives, and then made into prototypes for validation with user research.

The approach outlined here should be further detailed, refined, and validated in future research.

Figure 8. How the 3 themes correspond to basic psychological needs for relatedness, autonomy, and competence, in line with self-determination theory (SDT).

XSL•FO
RenderX

Figure 9. A proposed approach for designing future gaming-based positive psychology interventions (PPIs).



Pointers for Future Research

There is a range of possible strategies to choose from to work to improve an intervention and increase user motivation. In this specific study and for the purposes of this paper, we explored one possible strategy: to attempt to make the PPI more engaging, user-friendly, and relevant for young, unemployed adults by creating a gaming-based intervention. There are other alternative strategies that could increase relevance, appeal, and adaptation to the needs of the target audience. Kelders et al [24] suggest the use of design and persuasive design techniques, including reward, praise and reminders [107] as a tool to increase motivation and retention. Others [13] suggest the use of SDT [97], as we also found some support for this study. These strategies should be explored further in future studies.

Furthermore, the alignment between gameplay and intervention does not rely solely on the crafting of the game. Although our study grounded ideas on learning from past empirical research where PPIs had been applied to different contexts, there was a lack of clarity and theoretical grounding for the user experience in itself, including a clear definition of the learning [35] that should happen within the game design space. Incorporating learning theory, such as experiential learning [41], along with motivation theory, such as SDT [97] and persuasive design [108], as a more complete theoretical framework for the game designing process may provide a stronger direction to the conceptual work for the practitioners involved. Designing for

XSL•FO
RenderX

behavioral and mindset change is increasingly relevant for design research and professional design practice [109], and there seem to be several gaps in understanding for design researchers and design teams who find themselves grappling with psychological and behavioral theories to produce interventions to support problem-solving of societal problems, such as youth unemployment.

Limitations

In this study, relevant participants were involved in a design process to capture their experience with designs in-the-making and take feedback into consideration in the design of revisions. We consider such early involvement a strength of the study. However, it also holds limitations; as the results are drawn from user experiences with *prototype* PPIs, the study does not provide user experiences resulting from a completed and verified PPI. Although the knowledge gained through the different stages of the design process is of substantial value to this area of research, future work is needed on experiences with fully functional gameplay PPIs to validate the findings of this study and to measure engagement, effectiveness, and adherence to the intervention. In addition, considering the fact that the positive psychologies implemented in the game mechanics were, to some extent, unclear to the participants after exposure, we cannot draw any conclusions on the experience of these in themselves at this stage. However, this was also not the purpose of this study.

Another limitation is the choice and availability of participants in the study. The population of unemployed young adults is highly heterogeneous. With our recruitment strategy, we are aware that we do not cover the entire range of end users, especially because we in part relied on voluntary registration and on contact with specific public welfare systems. Nevertheless, we find the involved participants to be within the scope of the studied PPI, and their feedback, hence, is of substantial benefit in understanding how the PPI may be experienced by representatives of this target group.

The third limitation concerns the context of the usability testing. Being observed by another person influences behavior (eg, Hawthorne effect), and participants likely spent much more time considering the prototypes than they would normally have. However, this approach was chosen because our interventions were prototypes and had unfinished functionality, which required a moderator to “fill the gaps” [110,111]. A fully self-administered and unmoderated use of a gameplay PPI would be a natural next step in future research.

Finally, it is important to note that, although our exploratory approach to insight into user perceptions of a game-based intervention for this target group is an important starting point for this area of investigation, future research is needed to establish the knowledge base needed to reliably provide such interventions. As part of this, we envision future studies with larger sample sizes and established scales as part of randomized controlled trials to gain further knowledge of the effectiveness of game-based interventions for this group and a basis for improvements in intervention design.

Conclusions

The study contributes insights into key user perceptions of game-based interventions for unemployed young adults. The contribution has implications for future game-like intervention design for this purpose. Our principal contribution is to explore engagement through a PPI, designed as an interactive game. We have described the iterative process of the development of a 3D-game concept, VitaNova, and have explored participants’ thoughts and feedback on their experiences. Although the participants were positive about the general idea of a game targeted toward unemployed young people, we found tensions between a PPI and an exciting game play and 3 themes that pinpoint priorities for future gaming implementations. Our study shows that interactive game design could make interventions more entertaining and engaging but can easily come into conflict with or undermine the intervention. We recommend aligning the gameplay narrative, objectives, and mechanics with intervention content and objectives to create engaging, relevant, and effective gaming-based PPIs that promote a more productive view of the challenges experienced by the young and unemployed.

Acknowledgments

The authors would like to thank the researchers at the University of Stavanger who contributed to this study: Venke F Haaland, Espen Sagen, Mari Rege, Hilde Ness Sandvold, and Wenche Ten Velden Hegelstad. The authors would also like to thank the Norwegian Welfare Administration in Sandnes and Skole- og Jobbresept at the University Hospital of Stavanger, as well as all the participants who took part in the study and gave their feedback on the different prototypes. The project is funded by the Norwegian Research Council (grant number 296390).

Conflicts of Interest

None declared.

Multimedia Appendix 1
A presentation of the different game prototypes.
[PDF File (Adobe PDF File), 2796 KB - humanfactors_v11i1e44423_app1.pdf]

References

XSL•FO
RenderX

1. Youth and the labour market. Organisation for Economic Cooperation and Development. URL: https://www.oecd-ilibrary.org/education/youth-not-in-employment-education-or-training-neet/indicator/english_72d1033a-en [accessed 2024-01-05]
2. Borghans L, Duckworth AL, Heckman JJ, ter Weel B. The economics and psychology of personality traits. J Hum Res 2008;43(4):972-1059. [doi: 10.1353/jhr.2008.0017]
3. Heckman JJ, Stixrud J, Urzua S. The effects of cognitive and noncognitive abilities on labor market outcomes and social behavior. J Labor Econ 2006 Jul;24(3):411-482. [doi: 10.1086/504455]
4. Roberts BW, Kuncel NR, Shiner R, Caspi A, Goldberg LR. The power of personality: the comparative validity of personality traits, socioeconomic status, and cognitive ability for predicting important life outcomes. Perspect Psychol Sci 2007 Dec;2(4):313-345 [FREE Full text] [doi: 10.1111/j.1745-6916.2007.00047.x] [Medline: 26151971]
5. Bettinger E, Ludvigsen S, Rege M, Solli IF, Yeager D. Increasing perseverance in math: evidence from a field experiment in Norway. J Econ Behav Org 2018 Feb;146:1-15. [doi: 10.1016/j.jebo.2017.11.032]
6. Deming DJ. The growing importance of social skills in the labor market*. Q J Econ 2017;132(4):1593-1640. [doi: 10.1093/qje/qjx022]
7. McGee RE, Thompson NJ. Unemployment and depression among emerging adults in 12 states, behavioral risk factor surveillance system, 2010. Prev Chronic Dis 2015 Mar 19;12:E38 [FREE Full text] [doi: 10.5888/pcd12.140451] [Medline: 25789499]
8. Fyhn T, Radlick RL, Sveinsdottir V. Unge som står utenfor arbeid, opplæring og utdanning (NEET): en analyse av unge i NEET-kategorien. Norce. 2021. URL: https://www.ks.no/contentassets/25072fbf9d4d4577899c335e431117d7/Unge-utenfor-arbeid-opplering-og-utdanning.pdf [accessed 2023-08-02]
9. Investing in youth. Organisation for Economic Cooperation and Development. URL: https://www.oecd-ilibrary.org/social-issues-migration-health/investing-in-youth-norway_9789264283671-en [accessed 2024-01-05]
10. Bunting M, Moshuus G. Young peoples' own stories about dropping out in Norway: an indirect qualitative approach. Acta Didactica Norge 2017 Aug 25;11(2):3. [doi: 10.5617/adno.3182]
11. Moshuus GH, Ask TS, Bordevich K, Bunting M, Gulløy E, Halvorsen TA, et al. Inaktive ungdommer - en av vår tids største utfordringer. Sluttrapport (2015 – 2018). FARVE-Fondet, NAV, Kunnskapsavdelingen. 2019. URL: https://openarchive.usn.no/usn-xmlui/handle/11250/2687785 [accessed 2023-08-02]
12. Isherwood K. Examining the effect of positive psychology interventions to improve psychological wellbeing, workplace motivation and re-employment prospects in North Wales. Bangor University. 2019. URL: http://tinyurl.com/3zvm6vsx [accessed 2023-08-02]
13. van der Vaart L, Van den Broeck A. Self-determination and positive psychology interventions: an extension of the positive activity model in the context of unemployment. In: Theoretical Approaches to Multi-Cultural Positive Psychological Interventions. Cham, Switzerland: Springer; 2019.
14. Seligman ME, Csikszentmihalyi M. Positive psychology: an introduction. Am Psychol 2000;55(1):5-14. [doi: 10.1037/0003-066X.55.1.5]
15. Rege M, Hanselman P, Solli IF, Dweck CS, Ludvigsen S, Bettinger E, et al. How can we inspire nations of learners? An investigation of growth mindset and challenge-seeking in two countries. Am Psychol 2021 Jul;76(5):755-767 [FREE Full text] [doi: 10.1037/amp0000647] [Medline: 33180534]
16. Yeager DS, Hanselman P, Walton GM, Murray JS, Crosnoe R, Muller C, et al. A national experiment reveals where a growth mindset improves achievement. Nature 2019 Sep 7;573(7774):364-369 [FREE Full text] [doi: 10.1038/s41586-019-1466-y] [Medline: 31391586]
17. Dweck CS, Yeager DS. Mindsets: a view from two eras. Perspect Psychol Sci 2019 May 01;14(3):481-496 [FREE Full text] [doi: 10.1177/1745691618804166] [Medline: 30707853]
18. Paunesku D, Walton GM, Romero C, Smith EN, Yeager DS, Dweck CS. Mind-set interventions are a scalable treatment for academic underachievement. Psychol Sci 2015 Jun;26(6):784-793. [doi: 10.1177/0956797615571017] [Medline: 25862544]
19. Walton GM, Crum AJ. Handbook of Wise Interventions: How Social Psychology Can Help People Change. New York, NY: Guilford Publications; 2020.
20. Yeager DS, Romero C, Paunesku D, Hulleman CS, Schneider B, Hinojosa C, et al. Using design thinking to improve psychological interventions: the case of the growth mindset during the transition to high school. J Educ Psychol 2016 Apr;108(3):374-391 [FREE Full text] [doi: 10.1037/edu0000098] [Medline: 27524832]
21. Dambrun M, Dubuy AL. A positive psychology intervention among long‐term unemployed people and its effects on psychological distress and well‐being. J Employ Couns 2014 Jun 02;51(2):75-88. [doi: 10.1002/j.2161-1920.2014.00043.x]
22. Patel V, Saxena S, Lund C, Thornicroft G, Baingana F, Bolton P, et al. The Lancet Commission on global mental health and sustainable development. Lancet 2018 Oct 27;392(10157):1553-1598. [doi: 10.1016/S0140-6736(18)31612-X] [Medline: 30314863]
23. Christensen H, Reynolds J, Griffiths KM. The use of e-health applications for anxiety and depression in young people: challenges and solutions. Early Interv Psychiatry 2011 Feb;5 Suppl 1:58-62. [doi: 10.1111/j.1751-7893.2010.00242.x] [Medline: 21208393]

XSL•FO
RenderX

24. Kelders SM, Kok RN, Ossebaard HC, Van Gemert-Pijnen JE. Persuasive system design does matter: a systematic review of adherence to web-based interventions. J Med Internet Res 2012 Nov 14;14(6):e152 [FREE Full text] [doi: 10.2196/jmir.2104] [Medline: 23151820]
25. Lattie EG, Adkins EC, Winquist N, Stiles-Shields C, Wafford QE, Graham AK. Digital mental health interventions for depression, anxiety, and enhancement of psychological well-being among college students: systematic review. J Med Internet Res 2019 Jul 22;21(7):e12869 [FREE Full text] [doi: 10.2196/12869] [Medline: 31333198]
26. Michel T, Tachtler F, Slovak P, Fitzpatrick G. A review of youth mental health promotion apps towards their fit with youth media preferences. Pervasive Health Technol 2019 Feb 12;19(17):e2. [doi: 10.4108/eai.13-7-2018.161419]
27. Smyth J. Critically Engaged Learning: Connecting to Young Lives. Lausanne, Switzerland: Peter Lang Publishing Group; 2008.
28. Vygotsky LS, Vygotskiĭ LS, Cole M. Mind in Society: The Development of Higher Psychological Processes. Cambridge, MA: Harvard University Press; 1978.
29. Gee JP. What video games have to teach us about learning and literacy. Comput Entertain 2003 Oct 01;1(1):20. [doi: 10.1145/950566.950595]
30. 2022 Essential facts about the video game industry. Entertainment Software Association. URL: https://www.theesa.com/resource/2022-essential-facts-about-the-video-game-industry/ [accessed 2023-08-07]
31. Adachi PJ, Willoughby T. The link between playing video games and positive youth outcomes. Child Dev Perspect 2017 Mar 17;11(3):202-206. [doi: 10.1111/cdep.12232]
32. Ferrara J. Playful Design: Creating Game Experiences in Everyday Interfaces. Brooklyn, NY: Rosenfeld Media; 2012.
33. Anderson SP. Seductive Interaction Design: Creating Playful, Fun, and Effective User Experiences. Indianapolis, IN: New Riders; 2011.
34. Tan M, Hew KF. Incorporating meaningful gamification in a blended learning research methods class: examining student learning, engagement, and affective outcomes. Australas J Educ Technol 2016 Dec 01;32(5). [doi: 10.14742/ajet.2232]
35. Grund CK. How games and game elements facilitate learning and motivation: a literature review. In: Proceedings of the Informatik 2015. 2015 Presented at: Informatik 2015; September 28-October 2, 2015; Cottbus, Germany URL: https://dl.gi.de/handle/20.500.12116/2112
36. Cugelman B. Gamification: what it is and why it matters to digital health behavior change developers. JMIR Serious Games 2013 Dec 12;1(1):e3 [FREE Full text] [doi: 10.2196/games.3139] [Medline: 25658754]
37. Abt CC. Serious Games. Lanham, Maryland: University Press of America; 1987.
38. Young MF, Slota ST, Cutter A, Jalette GC, Mullin G, Lai B, et al. Our princess is in another castle: a review of trends in serious gaming for education. Rev Educ Res 2012 Mar;82(1):61-89. [doi: 10.3102/0034654312436980]
39. Djaouti D, Alvarez J, Jessel JP, Rampnoux O. Origins of serious games. In: Ma M, Oikonomou A, Jain L, editors. Serious Games and Edutainment Applications. London, UK: Springer; Oct 13, 2011.
40. Deterding S, Dixon D, Khaled R, Nacke L. From game design elements to gamefulness: defining "gamification". In: Proceedings of the 15th International Academic MindTrek Conference: Envisioning Future Media Environments. 2011 Presented at: MindTrek '11: Academic MindTrek 2011; September 28-30, 2011; Tampere, Finland. [doi: 10.1145/2181037.2181040]
41. Kolb DA. Experiential Learning: Experience as the Source of Learning and Development, Second Edition. Upper Saddle River, NJ: FT Press; 2014.
42. Sharlanova V. Experiential learning. Trakia J Sci 2004;2(4):36-39 [FREE Full text] [doi: 10.4135/9781412950596.n56]
43. Fleming T, Dixon R, Frampton C, Merry S. A pragmatic randomized controlled trial of computerized CBT (SPARX) for symptoms of depression among adolescents excluded from mainstream education. Behav Cogn Psychother 2012 Oct;40(5):529-541. [doi: 10.1017/S1352465811000695] [Medline: 22137185]
44. Burns JM, Webb M, Durkin LA, Hickie IB. Reach out central: a serious game designed to engage young men to improve mental health and wellbeing. Med J Aust 2010 Jun 07;192(S11):S27-S30. [doi: 10.5694/j.1326-5377.2010.tb03689.x] [Medline: 20528704]
45. Zepke N. Student engagement in neo-liberal times: what is missing? High Educ Res Dev 2017 Sep 04;37(2):433-446. [doi: 10.1080/07294360.2017.1370440]
46. Whitton N. Digital Games and Learning: Research and Theory. Milton Park, UK: Taylor & Francis; 2014.
47. Cheng VW, Davenport T, Johnson D, Vella K, Hickie IB. Gamification in apps and technologies for improving mental health and well-being: systematic review. JMIR Ment Health 2019 Jun 26;6(6):e13717 [FREE Full text] [doi: 10.2196/13717] [Medline: 31244479]
48. van der Bijl-Brouwer M, Dorst K. Advancing the strategic impact of human-centred design. Design Stud 2017 Nov;53:1-23. [doi: 10.1016/j.destud.2017.06.003]
49. The Field Guide to Human-centered Design: Design Kit. San Francisco, CA: IDEO; 2015.
50. Swann C. Action research and the practice of design. Des Issues 2002;18(1):49-61 [FREE Full text] [doi: 10.1162/07479360252756287]
51. Lewin K. A research approach to leadership problems. J Educ Sociol 1944 Mar;17(7):392-398. [doi: 10.2307/2262546]

XSL•FO
RenderX

52. Kock N. Action research. Interaction Design Foundation - IxDF. 2014. URL: http://tinyurl.com/yvxwcv8a [accessed 2023-08-08]
53. Stappers P, Giaccardi E. Research through design. Interaction Design Foundation - IxDF. 2014. URL: https://www.interaction-design.org/literature/book/the-encyclopedia-of-human-computer-interaction-2nd-ed/research-through-design [accessed 2023-08-08]
54. Frayling C. Research in art and design. Royal College of Art. 1994. URL: https://researchonline.rca.ac.uk/id/eprint/384 [accessed 2023-08-27]
55. Braun V, Clarke V. Thematic Analysis: A Practical Guide. Thousand Oaks, CA: SAGE Publications, Inc; 2021.
56. Tekinbas K, Zimmerman E. Rules of Play: Game Design Fundamentals. Cambridge, MA: MIT Press; 2003.
57. Cockton G. Usability evaluation. Interaction Design Foundation - IxDF. 2014. URL: https://www.interaction-design.org/literature/book/the-encyclopedia-of-human-computer-interaction-2nd-ed/usability-evaluation [accessed 2023-08-08]
58. Lewis JR. Usability: lessons learned … and yet to be learned. Int J Hum Comput Interact 2014 Jun 18;30(9):663-684. [doi: 10.1080/10447318.2014.930311]
59. Lewis JR. Sample sizes for usability tests: mostly math, not magic. Interactions 2006 Nov 01;13(6):29-33. [doi: 10.1145/1167948.1167973]
60. Gothelf J, Seiden J. Lean UX: Applying Lean Principles to Improve User Experience. Sebastopol, CA: O'Reilly Media; 2013.
61. Knapp J, Zeratsky J, Kowitz B. Sprint: How to Solve Big Problems and Test New Ideas in Just Five Days. New York, NY: Simon & Schuster; 2016.
62. Star SL. Chapter 2 - the structure of ill-structured solutions: boundary objects and heterogeneous distributed problem solving. In: Gasser L, Huhns MN, editors. Distributed Artificial Intelligence. Burlington, MA: Morgan Kaufmann; 1989:37-54.
63. Huvila I, Anderson TD, Jansen EH, McKenzie P, Worrall A. Boundary objects in information science. J Assoc Inf Sci Technol 2017 Jun 08;68(8):1807-1822. [doi: 10.1002/asi.23817]
64. Akkerman SF, Bakker A. Boundary crossing and boundary objects. Rev Educ Res 2011 Jun;81(2):132-169 [FREE Full text] [doi: 10.3102/0034654311404435]
65. Ericson KA, Simon HA. Protocol analysis. In: Protocol Analysis: Verbal Reports as Data Revised Edition. Cambridge, MA: MIT Press; 1993.
66. Mascherini M, Ledermaier S, Eurofound. Exploring the diversity of NEETs. Eurofound. 2016. URL: https://policycommons.net/artifacts/1840981/exploring-the-diversity-of-neets/2584049/ [accessed 2023-08-08]
67. Chen E, Leos C, Kowitt SD, Moracco KE. Enhancing community-based participatory research through human-centered design strategies. Health Promot Pract 2020 Jan 25;21(1):37-48. [doi: 10.1177/1524839919850557] [Medline: 31131633]
68. Nielsen J. How many test users in a usability study? Nielsen Norman Group. 2012 Jun 3. URL: https://www.nngroup.com/articles/how-many-test-users/ [accessed 2023-08-10]
69. Smith JA, Flowers P, Larkin M. Interpretative Phenomenological Analysis: Theory, Method and Research. Thousand Oaks, CA: SAGE Publications, Inc; 2009.
70. Patton MQ. Qualitative Evaluation and Research Methods, 2nd Edition. Thousand Oaks, CA: SAGE Publications, Inc; 1990.
71. Unge som registrerer seg hos NAV: 4 av 10 kommer raskt i jobb uten hjelp. NAV. 2021 Nov 30. URL: http://tinyurl.com/54dzea2z [accessed 2023-09-01]
72. Raaum O, Rogstad J, Røed K, Westlie L. Young and out: an application of a prospects-based concept of social exclusion. J Soc Econ 2009 Jan;38(1):173-187. [doi: 10.1016/j.socec.2008.08.003]
73. Horgen EH. Hvorfor ulike arbeidsledighetstall? Statistisk sentralbyrå. 2020 Jan 23. URL: https://www.ssb.no/arbeid-og-lonn/artikler-og-publikasjoner/hvorfor-ulike-arbeidsledighetstall--410344 [accessed 2023-08-10]
74. Dweck CS, Yeager DS. A growth mindset about intelligence. In: Walton GM, Crum AJ, editors. Handbook of Wise Interventions: How Social Psychology Can Help People Change. New York, NY: The Guilford Press; 2021.
75. Beck JS. Cognitive Behavior Therapy: Basics and Beyond Second Edition. New York, NY: Guilford Publications; 2011.
76. Yeager DS, Bryan CJ, Gross JJ, Murray JS, Krettek Cobb D, H F Santos P, et al. A synergistic mindsets intervention protects adolescents from stress. Nature 2022 Jul 06;607(7919):512-520 [FREE Full text] [doi: 10.1038/s41586-022-04907-7] [Medline: 35794485]
77. Chen DJ, Lim VK. Strength in adversity: the influence of psychological capital on job search. J Organ Behav 2012 Jun 25;33(6):811-839. [doi: 10.1002/job.1814]
78. Moorhouse A, Caltabiano ML. Resilience and unemployment: exploring risk and protective influences for the outcome variables of depression and assertive job searching. J Employ Couns 2011 Dec 23;44(3):115-125. [doi: 10.1002/j.2161-1920.2007.tb00030.x]
79. Bellotti F, Berta R, De Gloria A, Primavera L. Enhancing the educational value of video games. Comput Entertain 2009 Jun 24;7(2):1-18. [doi: 10.1145/1541895.1541903]
80. Aronson E. The power of self-persuasion. Am Psychol 1999 Nov;54(11):875-884 [FREE Full text] [doi: 10.1037/h0088188]
81. Braun V, Clarke V. Using thematic analysis in psychology. Qual Res Psychol 2006 Jan;3(2):77-101. [doi: 10.1191/1478088706qp063oa]

XSL•FO
RenderX

82. Braun V, Clarke V. Can I use TA? Should I use TA? Should I not use TA? Comparing reflexive thematic analysis and other pattern-based qualitative analytic approaches. Couns Psychother Res 2020 Oct 18;21(1):37-47. [doi: 10.1002/capr.12360]
83. Herrman H, Stewart DE, Diaz-Granados N, Berger EL, Jackson B, Yuen T. What is resilience? Can J Psychiatry 2011 May 01;56(5):258-265. [doi: 10.1177/070674371105600504] [Medline: 21586191]
84. Duckworth AL, Steen TA, Seligman ME. Positive psychology in clinical practice. Annu Rev Clin Psychol 2005 Apr;1(1):629-651. [doi: 10.1146/annurev.clinpsy.1.102803.144154] [Medline: 17716102]
85. Michel T, Tachtler F, Slovak P, Fitzpatrick G. Young people's attitude toward positive psychology interventions: thematic analysis. JMIR Hum Factors 2020 Nov 09;7(4):e21145 [FREE Full text] [doi: 10.2196/21145] [Medline: 33164908]
86. Torous J, Nicholas J, Larsen ME, Firth J, Christensen H. Clinical review of user engagement with mental health smartphone apps: evidence, theory and improvements. Evid Based Ment Health 2018 Aug 05;21(3):116-119 [FREE Full text] [doi: 10.1136/eb-2018-102891] [Medline: 29871870]
87. Hainey T. Moving digital games for learning forward. Horiz 2016 Feb 08;24(1):132-136. [doi: 10.1108/oth-08-2015-0037]
88. Fleming T, Merry S, Stasiak K, Hopkins S, Patolo T, Ruru S, et al. The importance of user segmentation for designing digital therapy for adolescent mental health: findings from scoping processes. JMIR Ment Health 2019 May 08;6(5):e12656 [FREE Full text] [doi: 10.2196/12656] [Medline: 31066705]
89. Bartle R. Hearts, clubs, diamonds, spades: players who suit muds. Players Who Suit MUDs. URL: https://mud.co.uk/richard/hcds.htm [accessed 2022-11-10]
90. Kort-Butler LA. The well-being of gamers, video game players, and non-players. Soc Sci J 2021 Apr 15:1-17. [doi: 10.1080/03623319.2021.1910774]
91. Williams D, Yee N, Caplan SE. Who plays, how much, and why? Debunking the stereotypical gamer profile. J Comput Mediat Commun 2008 Jul 1;13(4):993-1018. [doi: 10.1111/j.1083-6101.2008.00428.x]
92. Ryan RM, Rigby CS, Przybylski A. The motivational pull of video games: a self-determination theory approach. Motiv Emot 2006 Nov 29;30(4):344-360. [doi: 10.1007/s11031-006-9051-8]
93. Yee N. Motivations of play in MMORPGs. In: Proceedings of DiGRA 2005 Conference: Changing Views – Worlds in Play. 2005 Presented at: DiGRA 2005 Conference: Changing Views – Worlds in Play; June 16-20, 2005; Vancouver, BC URL: http://www.digra.org/dl/db/06276.26370.pdf
94. Deci EL, Koestner R, Ryan RM. A meta-analytic review of experiments examining the effects of extrinsic rewards on intrinsic motivation. Psychol Bull 1999 Nov;125(6):627-68; discussion 692. [doi: 10.1037/0033-2909.125.6.627] [Medline: 10589297]
95. Hekkert PP. Design aesthetics: principles of pleasure in design. Psychol Sci 2006;48(2):157-172 [FREE Full text]
96. Honea H, Horsky S. The power of plain: intensifying product experience with neutral aesthetic context. Mark Lett 2011 Oct 7;23(1):223-235. [doi: 10.1007/s11002-011-9149-y]
97. Deci EL, Ryan RM. Facilitating optimal motivation and psychological well-being across life's domains. Can Psychol 2008 Feb;49(1):14-23. [doi: 10.1037/0708-5591.49.1.14]
98. Baum A, Fleming R, Reddy DM. Unemployment stress: loss of control, reactance and learned helplessness. Soc Sci Med 1986 Jan;22(5):509-516. [doi: 10.1016/0277-9536(86)90016-x] [Medline: 3704688]
99. Bronfenbrenner U. The Ecology of Human Development: Experiments by Nature and Design. Cambridge, MA: Harvard University Press; 1979.
100. Lewin K. Field theory and experiment in social psychology: concepts and methods. Am J Sociol 1939 May;44(6):868-896. [doi: 10.1086/218177]
101. Walton GM, Yeager DS. Seed and soil: psychological affordances in contexts help to explain where wise interventions succeed or fail. Curr Dir Psychol Sci 2020 Jun 14;29(3):219-226 [FREE Full text] [doi: 10.1177/0963721420904453] [Medline: 32719574]
102. Shandley K, Austin D, Klein B, Kyrios M. An evaluation of 'Reach Out Central': an online gaming program for supporting the mental health of young people. Health Educ Res 2010 Aug;25(4):563-574. [doi: 10.1093/her/cyq002] [Medline: 20150530]
103. Ryan RM, Deci EL. Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. Am Psychol 2000;55(1):68-78. [doi: 10.1037//0003-066X.55.1.68]
104. Stenros J. In defence of a magic circle: the social and mental boundaries of play. In: Proceedings of 2012 International DiGRA Nordic Conference. 2012 Presented at: DiGRA Nordic '12; June 6-8, 2012; Tampere, Finland URL: http://www.digra.org/digital-library/publications/in-defence-of-a-magic-circle-the-social-and-mental-boundaries-of-play/
105. Baker C. Trying to design a truly entertaining game can defeat even a certified genius. Wired. 2008 Mar 24. URL: https://www.wired.com/2008/03/pl-games-25/ [accessed 2023-08-10]
106. Habit heroes. Wikipedia. URL: https://en.wikipedia.org/w/index.php?title=Habit_Heroes&oldid=1120538825 [accessed 2022-11-07]
107. Ludden GD, van Rompay TJ, Kelders SM, van Gemert-Pijnen JE. How to increase reach and adherence of web-based interventions: a design research viewpoint. J Med Internet Res 2015 Jul 10;17(7):e172 [FREE Full text] [doi: 10.2196/jmir.4201] [Medline: 26163456]

XSL•FO
RenderX

108. Fogg BJ. A behavior model for persuasive design. In: Proceedings of the 4th International Conference on Persuasive Technology. 2009 Presented at: Persuasive '09; April 26-29, 2009; Claremont, CA. [doi: 10.1145/1541948.1541999]
109. Cash PJ, Hartlev CG, Durazo CB. Behavioural design: a process for integrating behaviour change and design. Design Stud 2017 Jan;48:96-128. [doi: 10.1016/j.destud.2016.10.001]
110. Schade A. Remote usability tests: moderated and unmoderated. Nielsen Norman Group. 2013 Oct 12. URL: https://www.nngroup.com/articles/remote-usability-tests/ [accessed 2023-08-22]
111. Vasalou A, Ng B, Wiemer-Hastings P, Oshlyansky L. Human-moderated remote user testing: protocols and applications. In: Proceedings of the 8th ERCIM Workshop, User Interfaces for All. 2004 Presented at: 8th ERCIM Workshop, User Interfaces for All; June 28-29, 2004; Vienna, Austria URL: https://reed.cs.depaul.edu/peterh/papers/b2hd-VasalouERCIM2004.html

Abbreviations

CV: curriculum vitae
IPS: Individual Placement and Support
NAV: Norwegian Labour and Welfare Administration
NEET: not in education, employment, or training
PPI: positive psychology intervention
SDT: self-determination theory

Edited by A Kushniruk; submitted 18.11.22; peer-reviewed by S Azer, B Jeong, C Rodriguez-Paras; comments to author 07.06.23; revised version received 02.09.23; accepted 20.11.23; published 18.01.24.

Please cite as:
Straand IJ, Følstad A, Bjørnestad JR
Exploring a Gaming-Based Intervention for Unemployed Young Adults: Thematic Analysis
JMIR Hum Factors 2024;11:e44423
URL: https://humanfactors.jmir.org/2024/1/e44423
doi:10.2196/44423
PMID:38236624

©Ingjerd Jevnaker Straand, Asbjørn Følstad, Jone Ravndal Bjørnestad. Originally published in JMIR Human Factors (https://humanfactors.jmir.org), 18.01.2024. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on https://humanfactors.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Evaluating the User Experience of a Smartphone-Delivered Sexual Health Promotion Program for Older Adults in the Netherlands: Single-Arm Pilot Study

Ana Correia de Barros[1], PhD; Mariëtte Bergmans[2], BSc; Kreshnik Hasanaj[3], HS; Drianë Krasniqi[3], MD; Catarina Nóbrega[4], MSc; Bruna Carvalho Carneiro[4], MSc; Priscila A Vasconcelos[4], MSc; Ana Luísa Quinta-Gomes[4], PhD; Pedro J Nobre[4], PhD; Joana Couto da Silva[1], BSc; Cristina Mendes-Santos[1], PhD

[1]Fraunhofer Portugal AICOS, Porto, Portugal
[2]PCOB, Utrecht, Netherlands
[3]SPRIGS RND, Pristina, Kosovo
[4]Center for Psychology, Faculty of Psychology and Education Sciences, University of Porto, Porto, Portugal

Corresponding Author:
Ana Correia de Barros, PhD
Fraunhofer Portugal AICOS
Rua Alfredo Allen, 455
Porto, 4200-135
Portugal
Phone: 351 220430306
Email: ana.barros@fraunhofer.pt

Abstract

Background: Sexual health is an important component of quality of life in older adults. However, older adults often face barriers to attaining a fulfilling sexual life because of issues such as stigma, lack of information, or difficult access to adequate support.

Objective: We aimed to evaluate the user experience of a self-guided, smartphone-delivered program to promote sexual health among older adults.

Methods: The mobile app was made available to community-dwelling older adults in the Netherlands, who freely used the app for 8 weeks. User experience and its respective components were assessed using self-developed questionnaires, the System Usability Scale, and semistructured interviews. Quantitative and qualitative data were descriptively and thematically analyzed, respectively.

Results: In total, 15 participants (mean age 71.7, SD 9.5 years) completed the trial. Participants showed a neutral to positive stance regarding the mobile app's usefulness and ease of use. Usability was assessed as "Ok/Fair." The participants felt confident about using the mobile app. To increase user experience, participants offered suggestions to improve content and interaction, including access to specialized sexual health services.

Conclusions: The sexual health promotion program delivered through a smartphone in a self-guided mode was usable. Participants' perception is that improvements to user experience, namely in content and interaction, as well as connection to external services, will likely improve usefulness and acceptance.

(JMIR Hum Factors 2024;11:e56206) doi:10.2196/56206

KEYWORDS

internet interventions; mobile health; mHealth; older adults; sexual health; smartphone; user experience; pilot study; mobile phone

XSL•FO
RenderX

Introduction

Background

Sexual health is a component of general health [1] and quality of life in older age [2]. However, older age is also associated with barriers to a fulfilling sexual life [3-5]. Many older adults are sexually active [6] but are at a higher risk than the general population to present sexual difficulties and dysfunctions. Older women often report decreased libido or lack of vaginal lubrication, whereas erection issues, reduced sexual desire, or being unable to reach orgasm are difficulties regularly reported by men [7]. In health care services, sexual difficulties are often untreated [8] and aggravated by poor communication related to lack of appropriate and case-specific information, lack of training among clinicians, or negative social beliefs and societal stigma, which makes it difficult for both patients and clinicians to bring about the topic [9]. Therefore, identifying the means of circumventing societal stigma and providing timely and adequate support are 2 important courses of action to promote sexual health among older adults.

As the prevalence of smartphone ownership and access to the internet increase [10], there is an opportunity to use these technologies to deliver ubiquitous sexual health support in an inconspicuous manner, that is, one that does not overly expose support seekers to fear of social judgment. Smartphones, as they are intimate technologies that ubiquitously accompany their owners, seem to be an adequate means for the delivery of sexual health promotion programs. Although there is evidence of the efficacy of internet-based sexual health interventions for sexual dysfunction [11] or sexual health education [12], the literature is nonexistent on smartphone-based sexual health interventions targeting older adults [13].

Critical to the acceptance and adoption of such technologies is the user experience they provide [14,15]. Coined by Don Norman [16], the term "user experience" was used by the author to characterize all the sets of experiences a user has with a product throughout a user journey, from intention to use until postuse reflections [17]. Therefore, the concept goes beyond usability, defined by International Organization for Standardization as "the extent to which a system, product or service can be used by specified users to achieve specified goals with effectiveness, efficiency and satisfaction in a specified context of use" [18]. Designing positive user experiences with mobile digital technologies for older adult users has been a focus of many studies because the levels of engagement have been low, thus hampering the potential health benefits of such technologies [19]. Research has found that older adults' user experience with mobile digital health could be improved if the technology considered potential user sensorimotor and cognitive issues, users' motivation, and social support [19] as well as if it promoted more personalized experiences and trust [20]. Although there are general guidelines on designing for accessibility and inclusive design [21,22], best practices for designing digital technologies for sensitive topics such as sexuality and intimacy are lacking [23]. Understanding older adults' experiences with such technologies in the topic of sexual health is critical to improving their acceptability, usability, and adoption, so that they can deliver positive outcomes. However, no study has yet reported on older adults' user experiences with smartphone-delivered sexual health promotion programs.

To address these gaps, we have designed a smartphone-based sexual health promotion program [24] under a European project called Anathema (reference AAL-2020-7-133-CP). This program was made available to older adults in a longitudinal study during which we assessed the participants' user experience with the software. The findings contribute to the body of knowledge on older adults' preferences, use, and appropriation of digital technologies for sexual health and the design of smartphone-based sexual health promotion programs targeting this population.

Aim

The aim of this study was to evaluate the user experience of Anathema, a self-guided, smartphone-delivered program to promote sexual health among older adults.

Anathema Mobile App Overview

The mobile app used in this study was developed using a participatory design approach [25], which involved users from 3 European countries using the following methods: questionnaires, interviews, focus groups, usability tests, and co-design workshops [23,26].

The app is available for Android and iOS operating systems and contains a sexual health promotion program tailored to older adults. The program, which has an 8-week duration, is organized into 5 modules (which include chapters and subchapters):

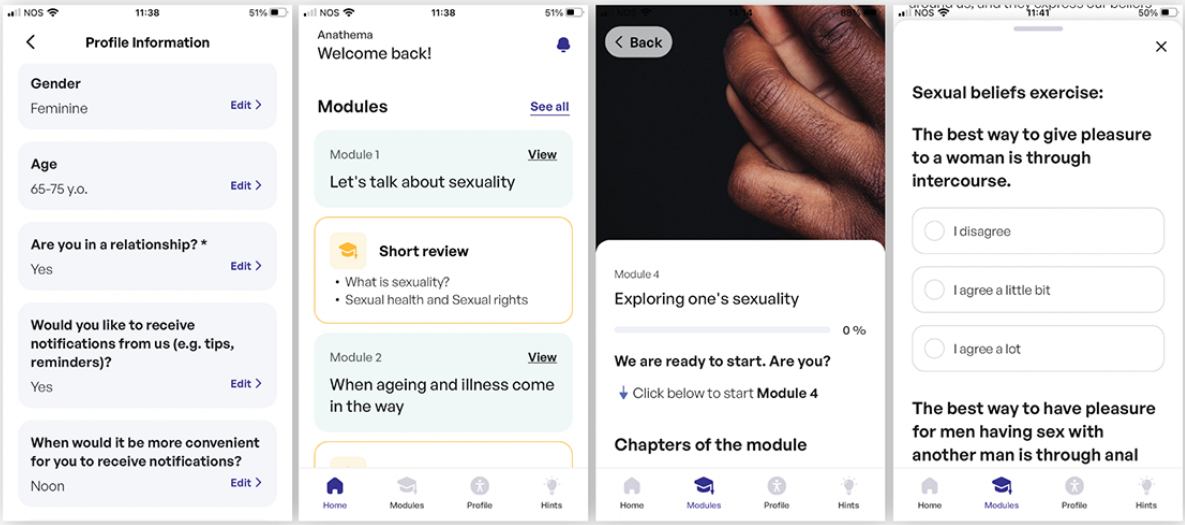
- *Module 1—Let's talk about sexuality (week 1)*: features information on male and female anatomies, sexual response, the importance of sexual pleasure, and sexual rights.
- *Module 2—When age and illness come in the way (week 2)*: addresses successful aging; the physiological, cognitive, and emotional changes in older age; and the main sexual problems and sexual dysfunction in older age.
- *Module 3—Emotional and physical intimacy (weeks 3-6)*: covers psychoeducation on the cognitive behavioral therapy model and the impact of sexual beliefs, thoughts, and emotions on sexuality. It includes exercises for cognitive restructuring, mindfulness, and communication skills training.
- *Module 4—Exploring one's sexuality (week 7)*: delivers information on sex aids and strategies to enhance sexual pleasure and satisfaction and includes sexual skills training and mindfulness exercises.
- *Module 5—Planning for a long-term fulfilling sex life (week 8)*: targets on relapse prevention with a focus on strategies to maintain progress and prevent setbacks. It also shares strategies to promote a healthy lifestyle and sexual health.

Each module is unlocked upon the completion of the previous module to ensure knowledge and skills acquisition. The chapters and subchapters are made of content in the form of text, images, and videos. The program also includes exercises such as written reflections or answers to multiple-choice questions using radio buttons (Figure 1). The app is available in English, European Portuguese, German, and Dutch languages.

XSL•FO
RenderX

The mobile app performs passive data collection through timestamp logs of interactions (eg, module completion date) as well as active data collection through logs of users' inputs on exercises. Another tool, Trial Monitor [27], fetches data from the database and shows visualizations thereof to the research or therapist teams.

Figure 1. Sample screenshots from the Anathema app (left to right): personal information, overview of modules, introduction to module, exercise.



Methods

Study Design

The study design was a single-arm pilot study with older adults (aged ≥55 years) testing the self-guided format of the sexual health promotion program in its Dutch version. The pilot study was conducted to assess user experience of the program. The content, structure, and format were also preliminarily evaluated toward the identification of improvements to the program and technological means of its delivery.

Inclusion and Exclusion Criteria

The inclusion criteria for participation in this trial were as follows: (1) being able to provide informed consent, (2) being aged ≥55 years, and (3) having digital skills and internet access. The exclusion criteria were as follows: (1) having a severe psychiatric disorder or alcohol or substance abuse; (2) taking medication that could interfere with sexual response; (3) having an uncontrolled medical condition that could interfere with sexual health; and (4) currently being on psychotherapy for sexual or intimate problems or for other psychological problems or current participation in another intervention study or clinical trial (or both).

Study Procedures

In a previous phase of this research, 1119 older adults, recruited through the contact list of the Dutch senior organization Katholieke Bond van Ouderen - Protestant Christelijke Ouderen Bond (KBO-PCOB), answered a questionnaire on unmet sexual needs [26]. In this questionnaire, the respondents were asked to indicate whether they would be available for future research within the same research project. Respondents who gave a positive reply were regularly invited to participate in user research activities throughout the research project [23], including the pilot study described in this paper. The majority of this subsample (N=346) were men (69.4%), had a high education level (53.2%), and were retired (89.9%). For the pilot study, further potential participants were contacted via other KBO-PCOB channels, including KBO-PCOB's employees.

Upon signing the informed consent form, participants were asked to complete a web-based screening questionnaire. If deemed eligible to participate in the study, the participants were asked to answer a web-based sociodemographic questionnaire. After completing the questionnaire, participants were provided access to the Anathema app and were prompted to complete the program in 8 weeks.

Once they had completed the 5 modules in the app, participants were asked to fill in a web-based, self-developed user experience questionnaire (Multimedia Appendix 1), which also included the System Usability Scale (SUS). Participants were then invited to participate in a semistructured debriefing interview about (Multimedia Appendix 2) their experiences with the program.

Metrics and Data Analyses

The main outcome of the study was user experience, which included dimensions of usefulness and usability. User experience was assessed after the intervention with a self-developed multiple-choice list of characteristics (answer options: *accessible*, *arousing curiosity*, *attractive*, *boring*, *elegant*, *fascinating*, *helpful*, *instructive*, *meets expectations*, and *strenuous*), a question on free grading of the app from 1 to 10, with 10 being the highest grade, a Net Promoter Score question (answer options: *Yes*/*No*/*Don't know*), and a semistructured debrief interview with questions addressing usefulness, usability, feasibility, clinical aspects, and implementation (Multimedia Appendix 2). Perceived usefulness was assessed using a self-developed 5-point Likert scale assessing the program in general, each module, and exercises. Usability was assessed using a self-developed 5-point Likert scale on perceived ease of use and perceived readability, as well as with the Dutch

XSL•FO
RenderX

version of the SUS [28]. Assessment of the self-perceived contribution of the program to changes in satisfaction and pleasure in sex life was also performed postintervention with a single-item question (4-item descriptive rating scale).

To characterize the study sample, sociodemographic variables were collected using a self-developed questionnaire assessing age, education, professional status, gender, sexual orientation, marital status, current sexual partnership status, satisfaction with current sex life (5-point Likert scale), self-rated quality of life, and degree of satisfaction with their own health (based on items 1 and 2 from World Health Organization Quality of Life Brief Version [29]).

The interviews were audio-recorded and partially transcribed for relevant content. The transcriptions, written in Dutch, were then translated into English by a native Dutch speaker (MB) for analysis by a non-Dutch speaker (ACB). The questionnaire and the interview data were analyzed descriptively and thematically, respectively.

Ethical Considerations

The study was approved by the Ethics Committee of the Faculty of Psychology and Educational Sciences, University of Porto (reference 2022/01-05b). All potential participants were informed about the study objectives and procedures. The participants who agreed to participate signed the informed consent form. There was no compensation or payment offered to the participants.

Results

Participants

A total of 400 participants were approached to participate in this study. Most participants did not provide a reason for declining or not answering the invitation. Among those who did (n=47), the reasons given were that participants were no longer interested (n=15), considered the pilot required too much commitment or effort (n=12), felt uncomfortable with the topic (n=9), considered they did not meet the criteria (n=5), or had a malfunctioning email (n=5). We also received information that one person had died.

In total, 23 participants agreed to participate and completed a web-based screening questionnaire to confirm the eligibility criteria. All participants were deemed eligible and were given access to Anathema after answering a sociodemographic questionnaire.

A total of 8 participants dropped out of the study. Of them, 4 participants did not provide any reasons for abandoning the study. Those who did shared the following reasons: discontinued access to the internet (n=1), dissatisfaction with the fact that future content modules were locked (n=1), inability to install and open the app (n=1), and lost motivation to use the app (n=1). A total of 15 participants used the Anathema app, having completed all the modules and completed the final questionnaire on user experience and usability. In total, 8 participants agreed to participate in a debriefing interview.

The 15 participants who used the app and answered the final questionnaire were 7 cisgender women and 8 cisgender men aged between 56 and 85 years (mean 68.3, SD 9.5 years). Most (n=12) were retired, and most (n=10) had completed higher professional education. Overall, 6 participants were married, 4 were single, 3 were cohabiting, and 2 were widowed (Table 1).

Table 1. Sociodemographic characteristics of the sample (N=15).

Characteristics	Values
Gender, n (%)	
Female	7 (47)
Male	8 (53)
Marital status, n (%)	
Single	4 (27)
Cohabiting	3 (20)
Married	6 (40)
Widowed	2 (13)
Professional status, n (%)	
Employed	3 (20)
Retired	12 (80)
Education, n (%)	
Secondary professional education	2 (13)
Higher professional education	10 (67)
University or scientific training	3 (20)
Age (years), mean (SD; range)	71.7 (9.5; 56-85)

XSL•FO
RenderX

Most of the 15 participants were exclusively heterosexual (n=12), most had sex with a partner in the context of an exclusive relationship with that person (n=11), and the level of sexual satisfaction was heterogeneously distributed, as shown, together with complete sexual characteristics (Table 2). The sample comprised participants who tended to positively rate their quality of life and health (Table 3).

Table 2. Sexual characteristics of the sample (N=15).

Characteristics	Baseline, n (%)
Sexual orientation or preference	
Exclusively heterosexual	12 (80)
Mainly heterosexual	2 (13)
Exclusively homosexual	1 (7)
Current sexual partners	
Sex with a partner, in the context of my exclusive relationship with him or her	11 (73)
Casual sex with a partner	1 (7)
No sexual partner	3 (20)
Satisfaction with current sex life	
Very satisfied	3 (20)
Satisfied	5 (33)
Neither satisfied nor dissatisfied	4 (27)
Dissatisfied	3 (20)

Table 3. Perceived quality of life and health satisfaction (N=15).

	Baseline, n (%)	After the test, n (%)
Rating of quality of life[a]		
Very good	8 (53)	9 (60)
Fairly good	7 (47)	5 (33)
Neither good nor bad	__[b]	1 (7)
Satisfaction with health[c]		
Very satisfied	8 (53)	7 (47)
Satisfied	7 (47)	7 (47)
Neither satisfied nor dissatisfied	—	1 (7)

[a]Original wording: How would you rate your quality of life? Responses were rated on a 5-point Likert scale: 1=very bad to 5=very good.

[b]Not available.

[c]Original wording: How satisfied are you with your health? Responses were rated on a 5-point Likert scale: 1=very dissatisfied to 5=very satisfied.

User Experience

In this section, we present the quantitative and qualitative results of the participants' user experience (Table 4). As we do so, we provide interpretations of the results mostly because of the interpretation required by the analysis of the interview data. Therefore, we discuss some of the results as we present them.

XSL•FO
RenderX

Table 4. Results of the user experience questionnaire (N=15).

	Values
Would recommend Anathema to friends or family (net promoter score)[a], n (%)	
Yes	6 (40)
No	6 (40)
Doesn't know	3 (20)
Perceived usefulness of app[b], n (%)	
Very useful	2 (13)
Useful	6 (40)
Neither useful nor useless	6 (40)
Useless	0 (0)
Extremely useless	1 (7)
Perceived usefulness of exercises[c], n (%)	
Very useful	0 (0)
Useful	7 (47)
Neither useful nor useless	3 (20)
Useless	3 (20)
Extremely useless	2 (13)
Perceived ease of use[d], n (%)	
Very easy	1 (7)
Easy	7 (47)
Neither easy nor difficult	5 (33)
Difficult	2 (13)
Readability[e], n (%)	
Very easy	1 (7)
Easy	7 (47)
Neither easy nor difficult	5 (33)
Difficult	2 (13)
System Usability Scale score, mean (SD; range)	56.3 (19.1; 20-85)
Score (1-10) given to Anathema app, mean (SD)	6.5 (1.8; 2-9)
Perceived impact of Anathema app in satisfaction and pleasure[f], n (%)	
Positive impact	4 (27)
No change	7 (47)
Negative impact	1 (7)
Doesn't know	3 (20)

[a]Original wording: Would you recommend the Anathema app to friends and/or family members?

[b]Original wording: How useful do you think the Anathema app is for older adults? Responses were rated on a 5-point Likert scale: 1=extremely useless to 5=very useful.

[c]Original wording: How useful did you find the (writing) exercises you were offered? Responses were rated on a 5-point Likert scale: 1=extremely useless to 5=very useful.

[d]Original wording: How easy was it for you to use the Anathema app without any help from others? Responses were rated on a 5-point Likert scale: 1=extremely difficult to 5=very easy.

[e]Original wording: How readable did you find the content of the Anathema app? Responses were rated on a 5-point Likert scale: 1=extremely difficult to 5=very easy.

[f]Original wording: Do you have the impression that the Anathema app can help you change satisfaction and pleasure in your sex life? Rated using a

XSL•FO
RenderX

descriptive scale: Don't know; No, no change; Yes, namely less satisfying and fun; Yes, namely more satisfying and fun.

Most participants showed a neutral to positive stance toward the app regarding its *usefulness*. There are some nuances when analyzing the perceived usefulness per module, as illustrated in Figure 2. Modules 2 and 3 had slightly more polarized responses. Modules 1 and 2 were found to be "very useful" for more participants, likely because of the reasons given in the interviews: participants learned new concepts, learned to understand what is normal in aging ("I end up thinking about the part about body ageing. That's reliable information that I can't easily get anywhere else today" [P03]), were made to rethink the way in which they faced sexuality, and also learned about the genitalia of other sexes:

> *Nice to read some details about genitals[...] also from the opposite sex, how something works.* [P11]
>
> *Enlightening. I did benefit from seeing what a prostate looked like.* [P04]

Figure 2. Visualization of perceived usefulness by module.

Module 1	Module 2	Module 3	Module 4	Module 5
Extremely useless	Extremely useless	Extremely useless	Extremely useless	Extremely useless
Useless	Useless	Useless	Neither useful nor useless	Neither useful nor useless
Neither useful nor useless	Useless	Useless	Neither useful nor useless	Neither useful nor useless
Neither useful nor useless	Neither useful nor useless	Neither useful nor useless	Neither useful nor useless	Neither useful nor useless
Neither useful nor useless	Neither useful nor useless	Neither useful nor useless	Neither useful nor useless	Neither useful nor useless
Neither useful nor useless	Neither useful nor useless	Neither useful nor useless	Neither useful nor useless	Neither useful nor useless
Neither useful nor useless	Neither useful nor useless	Neither useful nor useless	Useful	Useful
Neither useful nor useless	Useful	Neither useful nor useless	Useful	Useful
Useful	Useful	Useful	Useful	Useful
Useful	Useful	Useful	Useful	Useful
Useful	Useful	Useful	Useful	Useful
Very useful	Very useful	Useful	Useful	Useful
Very useful	Very useful	Very useful	Very useful	Useful
Very useful	Very useful	Very useful	Very useful	Useful
Very useful	Very useful	Very useful	Very useful	Very useful

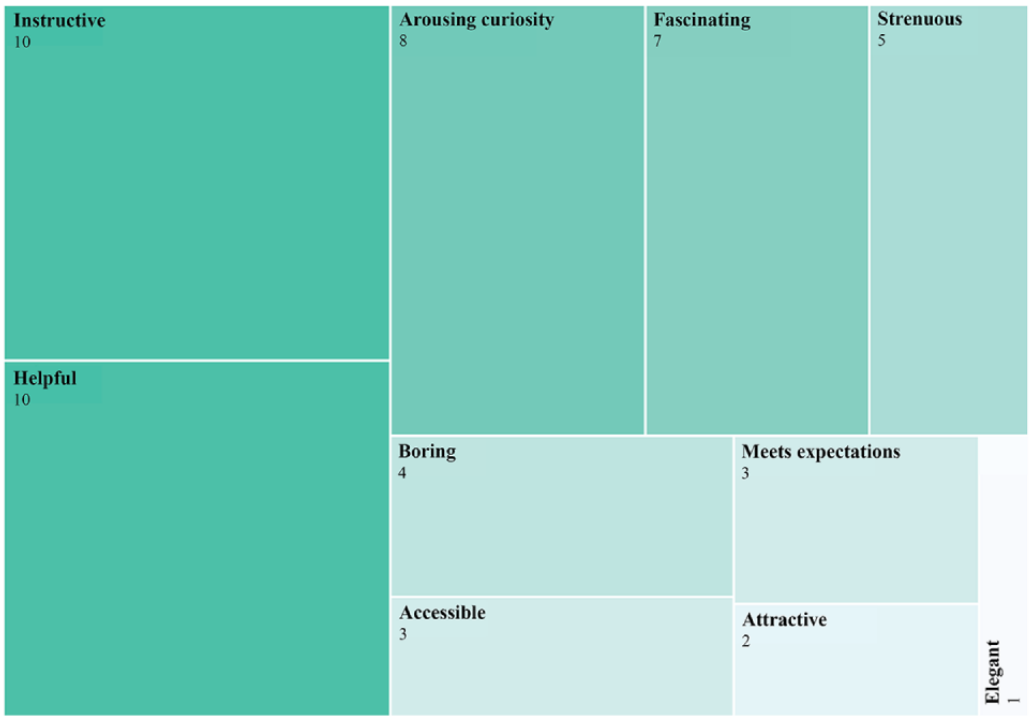
Other highlighted learning points from the program are the importance of communication and the fact that sexuality does not need to be equated with penetration. Something some participants missed was the possibility to ask the app questions about their specific problems, ask questions anonymously, or to be able to search for certain themes that could be of more interest to them.

Seven participants evaluated the *exercises* as useful, whereas the other 8 found them neutral (n=3), useless (n=3), or extremely useless (n=2). Crossing these results with the information provided in the interviews, one can infer that there were 2 aspects that hindered the experience with the exercises: on the one hand, participants struggled with long text input on their smartphone keyboards; on the other hand, for this group, the feeling of being "schooled" by the app was not equated with positive emotions, thus negatively impacting the experience. Finally, in the interviews, participants revealed that some exercises helped them think of sexuality in a different way, which they experienced as being positive.

When asked to attribute *characteristics* to the *Anathema app*, most participants selected a set of descriptors displayed in Figure 3, but the number of choices varied from a single adjective to 6 adjectives. Most of the qualifiers have positive valence, with the exception of "boring" and "strenuous," with 4 and 5 mentions, respectively. In line with the data collected through the interviews, the participants perceived that they had learned from the app. However, only 5 participants assessed the app as having the potential to help change their sexual satisfaction and pleasure.

XSL•FO
RenderX

Figure 3. Visualization of qualifiers attributed to the Anathema app and how often they were attributed.



The interviews also revealed that participants appreciated the app aesthetically, which connects to the descriptors that were chosen, as well as the tone of voice that was adopted for the content, which, in some cases, helped them deal with a sensitive topic:

> *I admire that this can be done in an app. Good looking and doesn't scare someone. I managed to deal with such a sensitive topic. [It's] friendly and nicely constructed.* [P22]

For 2 participants, the communication style options were not the most appropriate, for example, when showing an animated video of an anthropomorphized clitoris. Although the photographs were selected based on a survey conducted by the research team about the characteristics of photos that were appreciated by Dutch older adults, 2 interviewees did not find them totally appropriate, for example, some having a comical or childish tone, representing too young people, or not representing enough diversity.

Taking the *net promotor score* as an indicator of satisfaction, we can see that opinions were divided. Three participants did not know whether they would recommend the app to friends or family, whereas the remaining 12 participants were equally divided between wanting to recommend and not wanting to do so. In the debriefing section of the interviews, participants who were not certain whether to recommend Anathema expanded on this. They explained that they think the app has potential but that it needs certain improvements, as described earlier, for them to confidently recommend it to others.

The average SUS score, which measures *usability*, stood at 56.3, which, according to the scoring standards, corresponds to an assessment of "OK/Fair" [30]. Based on the averages per item, we can see that participants tend to think that they do not need help in using the system, although usability is not perceived to be at the excellent level. The level of confidence felt by participants while operating the app was high. Participants generally showed a neutral to positive stance toward the app regarding its *ease of use* and its *readability*. Although most people did not experience trouble reading because of font size or contrast, this was an issue for one of the participants who dropped out:

> *I also found the fine print difficult. They are clear but with deteriorating eyes good reading requires more effort.* [P04]

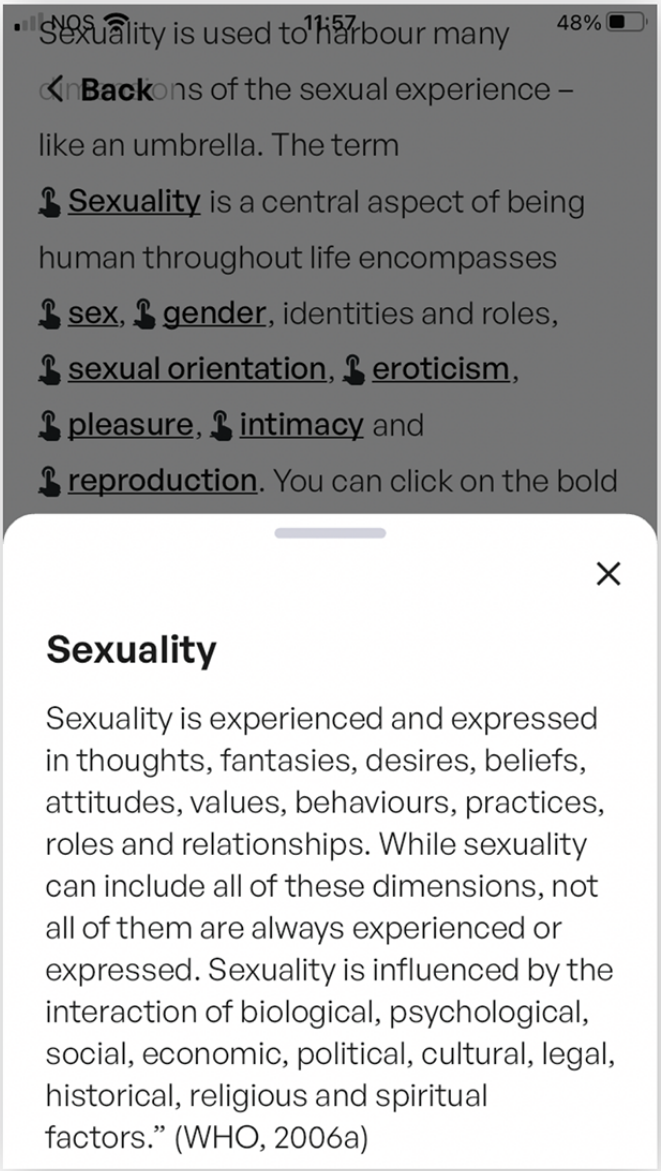
The interviews revealed that the app worked well on participants' phones and that they found it very convenient. However, participants often wished that the app would also be easy to use on a tablet device or desktop:

> *Excellent [the experience of using the app on the phone]. Preferably on an iPad, because of the larger screen. On the phone it worked. The smaller keyboard asked more caution, but [it] went fine.* [P11]

With the exception of 2 participants, who suggested direct speech, easier wording, and shorter sentences, interviewees found the wording easy to understand. In total, 2 participants reflected on whether the scientific explanation should be highlighted as is (Figure 4), for instance, on starting the first module with the definition of sexuality or whether it should be made more digestible to engage readers.

XSL•FO
RenderX

Figure 4. Screenshot from the definition of sexuality, which appears in a slide-up pop-up when the user taps the word "Sexuality" shown in the gray-faded part of the image. WHO: World Health Organization.



Another aspect of readability that was touched on was finding one's place in the content structure. For 2 participants, it was hard to understand at which stage they were in navigating the app, originating the feeling of being lost: "In a book you can browse through that and then you see where you are. In the app this overview is not so clear" (P06). For those who felt lost, as well as for one participant who would like to revisit specific parts of the content, a possible solution was provided by one participant who said that they missed a way to bookmark "Favorites."

The participants *appropriated* the app in different ways. There were reports of people using the app only randomly when they found the time, defining a fixed schedule (eg, evening, late at night), or defining a place to use the app (eg, kitchen, home). Common to the participants was the need to use the app alone and undisturbed.

One participant asked the partner to also go through the app, but they did not want it because it was a taboo topic. Three mentioned how they talked to friends or their partners later about the app and what they had learned, for example:

> *Through the app I can easily talk with my partner about sexuality. The participation, together with my partner, in the previous workshops for Anathema, also contributed to this. The openness of other participants was a good example for me.* [P04]

While 5 participants stated that there would be no place or time when they felt uncomfortable using the app, the remaining 3 gave some indications thereof. For these participants, it would be important to use the app alone and in a private place. Participants also reflected on how they would like to discuss what they were learning in the app with others but found it stigmatizing:

> *When I talk to friends about food, for example, all the experiences can be discussed. Apparently, that is not possible when talking about sex.* [P20]

XSL•FO
RenderX

> *When I try to discuss with seniors of an association with a Catholic background that I am participating in this project, the reaction is that does not suit our people.* [P18]

Half of the participants found an *8-week period* to be too long, whereas the other half felt it was an acceptable or good duration. However, participants struggled with the idea of this being time bound in some way because they could not understand why this specific duration was chosen. In 1 case, the participant felt that this indication of duration could send the wrong message: "I have to be ready in 8 weeks" (P03).

Being presented with *content that had a specific reading order* was cumbersome to some participants. This was because, on the one hand, they could feel schooled, and, on the other hand, they did not want to feel that they were losing time in content that they were not interested in. One participant shared their technique for when something like this happened: they just scrolled the content very quickly to get to the bottom and move to the next chapter. Despite negative comments about the locked content (eg, "I wanted to look at a topic in Module 4. But didn't do that out of irritation at the locks in the extended Module 3" [P04]), participants generally agreed that the content is well structured, being easy to follow. Mindfulness is something that some interviewees found unnecessary. On the other hand, some interviewees would expect to read more about love and affection. There were also other suggestions of curated lists of contacts for further support (eg, participants stated they would like to be able to ask questions to therapists over email) and fitness exercises (eg, pelvic floor muscle exercises).

Although 9 of 15 participants in the questionnaire assessed module 5 as useful or very useful, the interviews revealed a slightly different picture. The interviewees had mixed opinions regarding the usefulness of the last module. With the exception of 1 person, those who found it useful as a recap also reflected on the possibilities of *coupling the app* with curated contacts to therapists to continue exploring the topic or to find tailored help to a specific issue. One interviewee thought about accompanying the app with television or radio shows, stating that this was the reason why they bought a book on sexuality. Another possible extension would be a course, workshops, or group activities that would let people discuss and further explore what they had learned and experienced:

> *In addition to using the app, it could be interesting to be in a discussion group with other couples as a couple. That could help improve communication about sex. The app provides plenty of conversation material for that.* [P11]

Interviewees had mixed opinions regarding *whether the app should be paid.* On the one hand, participants shared that they are not used to paying for apps, but on the other hand, they recognized that they might pay for extra services (eg, consultations) and that free apps do not have as much credibility. Credibility is something that participants cling to when reflecting with the interviewer about how to make the app available to more people. Participants concluded that the app could be credibly made available through medical doctors, therapists, or reliable associations. Although this was not asked, participants also shared ideas on how to raise awareness about the Anathema app, for example, through advertisement, television or radio shows, or leaflets.

Discussion

Principal Findings

The pilot study conducted in the Netherlands with a group of 15 community-dwelling older adults was a novel study in the field of mobile health apps in sexual health. Although the dropout rate was high (65%), no participants were lost to follow-up or nonuse cases, that is, participants answering the questionnaires without having used the mobile app until the end. We found that the app was usable, that participants showed high levels of self-confidence in using it, that the smartphone can be a useful and private way to have access to reliable sexual health information, that participants foresee how extra services could help tailor the program to their specific needs, and that certain improvements in content and in interaction are likely to increase user experience for this smartphone-delivered sexual health promotion program.

As with other studies in the literature [20], the user experience was negatively affected by a lack of social support for users' specific issues. In the interviews, participants gave examples of further content on love and affection, a curated list of resources and fitness exercises that they would like to see, and options to search through the content to get the information they were looking for. The lack of social support, ranging from relatives to professionals, also seems to have negatively affected participants' user experiences. In their systematic review, van Acker et al [19] noted how social support (ranging from relatives to professionals) was an important factor in user experience. In our study, with the exception of 1 participant who could not convince their sexual partner to also use the app, there were no reports of available or lacking support from relatives, but participants specifically mentioned that professional support would be useful in addition to the existing offer. A nuance with relation to the literature [19] is that participants in our study did not require much professional support to interact with the program, but rather as an extension to it, often to attain the personalization requirement we have just described earlier. Furthermore, the participants struggled with the locked content. Although the tunneling technique has been used to increase engagement with intervention or technology, in our study, it did not seem to have this effect. This is similar to recent findings with an intervention for a younger generation [31].

As noted in the literature [20], trust is also an important dimension in user experience. Although not directly asked about it, our interviewees alluded to the element of credibility regarding willingness to pay, which was considered by Hurmuz et al [20] as a metric of user experience. For the participants in our sample, the channel via which they access the app is an important factor at the time of choosing whether to use and ultimately pay for the app.

As measured by the SUS instrument, self-confidence among the participants in our study was high. This might also have been influenced by the level of education and digital literacy of

XSL•FO
RenderX

the sample. The level of education might also explain why the participants often alluded to the experience of "being schooled" as a negative valence. Although the tone of voice for the program regarding visual and written content was co-designed [23], it might not have been implemented properly to eliminate this negative experience. This aspect is further discussed in the "Limitations" section below. On the other hand, some users also reacted negatively to content that seemed "too scientific," and some commented that some terms might not be easy to understand for the wider population. This is at odds with the higher educational level of this sample, but the explanation for the dislike might be related not to the understandability of the content but rather to a kind of experience that users expect when they are using an app that is related to sexuality.

The participants stated that the topic of sexuality was not embarrassing. However, there were some accounts of users requiring privacy when going through the content, one user whose partner did not want to use the app because of the topic, or users commenting on how they did not feel free or at ease discussing the topic with their peers. Therefore, the topic of taboo still requires further research in terms of how much of a barrier it is to accept and use technology around this topic. Participants' statements in the interviews suggested that a smartphone-based intervention can bring the advantages of ubiquity, intimacy, and anonymity to an intervention that is likely to elicit stigma in some contexts. The program itself was regarded as a trustworthy source of information that participants think is difficult to find on this topic. On the other hand, it could be coupled with more targeted personal services for users who would like to interact with therapists or even join groups willing to openly discuss topics of sexual health. Future research should study the provision of such discussion groups either in person or through moderated and anonymous forums inside the app.

Our study included participants interested in sexual health. In any case, even within our small sample, we witnessed a wide spectrum regarding taboo. For instance, some participants felt blurring genitalia photographs by default with overlaid text: "Sensitive content. Click to view" was condescending, whereas others felt that suggesting exercises for sexual pleasure was going too far. As with other types of apps targeting older adult users, our study saw a large heterogeneity in user preferences. Even if resources are allowed for the software development team to implement ultrapersonalization, we could place a large burden on users upon onboarding to set up preferences, which, in itself, would have a negative effect on technology acceptance. One way of addressing this could be to create certain user profiles and adapt scaffolding techniques that have been used for usability [32] for the purpose of conspicuousness degrees. Future research should work on this balance between a certain level of tailoring to one's needs and preferences, with time invested in customizing the app.

Strengths

This was the first study to evaluate the user experience of a self-guided, smartphone-delivered program to promote sexual health among older adults. The mixed methods approach was a strength of this study in the sense that it provided a rich description of participants' experiences with the app and the program. Without the interviews, we would hardly have had such detailed information that would indicate how to improve the app and the program, as well as a first understanding of how participants appropriated the app.

Our study did not aim at generalizability but rather at an in-depth understanding of user experience, which justified the emphasis on the qualitative data. Through this approach, we derived actionable insights to improve the content, structure, and format of the program.

Although our study was conducted with a small and specific sample of older adults in the Netherlands, the methodology we used allowed us to unveil nuances that can be useful for researchers to consider when implementing smartphone-based programs for sexual health in different populations: the relevance of social support, the credibility of the program, the opportunities that smartphone-based interventions may bring to sexual health interventions in terms of privacy or convenience, and the variability among program users about what might be considered a taboo and how this might impact users' preferences, practices, and attitudes toward the programs.

Limitations

As we conducted a user experience pilot study to obtain in-depth feedback, the results might not reflect the characteristics of the older adult population in the Netherlands. Although further research is needed to reach generalizability, this study constitutes a stepping stone in this journey.

The sample characteristics in our study are its greatest limitation. Only one-fifth of our participants were dissatisfied with the current state of their sex lives, and most considered themselves to be in fairly good or good health, which may not be representative of the older adult population. These characteristics may have biased how participants responded to a sexual health promotion program tailored to help users identify and cope with issues related to their sexual health. Our sample also comprised participants with a high level of education. This might explain why some participants felt schooled, as they were already in possession of information that was provided by the program. As participants have suggested, for a future pilot study, it would be advisable to increase the number and type of channels used for dissemination and recruitment, such as the mainstream media. This would help increase the visibility of Anathema and reduce, if not altogether, prevent, selection bias.

The features implemented on the app responded as much as possible to the user research requirements, but this was not always possible or perhaps implemented at its best. In some cases, there were technical limitations that did not allow their implementation. For instance, the app began to be implemented as web based so that it would also run on desktop browsers if participants preferred, but the identification of a problem in a technical component ahead of the implementation process forced the software development team to develop natively for Android and iOS.

We expected this lack of flexibility in the device type to be a negative aspect for some participants. On the other hand, at least once, the preferences collected from participants in user research studies preceding the pilot study were not aligned with the

XSL•FO
RenderX

preferences of the pilot study sample. We describe 2 instances of this problem.

The first example relates to the choice of imagery. To select the photos for the app, we conducted a survey with 111 older adults in the Netherlands, in which we showed 10 different pictures and asked participants to rate the pictures, select their favorites, and justify their choices. The survey revealed that participants preferred uplifting, cheerful, and romantic images of participants who were not young but also not too old. The interviews in the pilot study revealed that, for some participants, these images were not appropriate.

Another example was the language used: a series of tests on the preferred tone of voice were used to create the original content in English [23]. The content was translated into Dutch, which went through content reviews from native speakers with experience in older adult care. Nevertheless, for some participants in the sample, the language was described as "too scientific." It is also possible that the research team was not able to correctly implement the insights from the user research phase, thus causing a mismatch between the users' expectations and the implemented app. Further research should revise the feedback from the user research phases and cross it with the results from the pilot study to understand where the app can be improved to meet users' expectations.

Further research should also focus on interaction and content issues to improve current mobile apps toward improving user experience. In particular, there is a need to understand how to balance the quantity and type of content with an engaging user experience. Once an improvement in user experience has been noted through further formative testing, the pilot should be repeated. As there was a mix of negative and positive comments provided by the participants and because the results from the SUS score are aligned with the comments from the interviews, we do not think that social desirability influenced participants' answers. However, as social desirability plays an important role in sex research surveys, a future pilot study could include a questionnaire (eg, [33]) to control for this effect. Further research should focus on a larger and more diverse sample regarding sexual satisfaction, health status, and literacy level.

Future pilots should include study designs that enable the collection of fine-grained data about the user experience combined with an assessment of the program's efficacy in improving sexual health so that the aspects of appropriation and how the app fits into participants' practices could be better understood and, in turn, inform strategies to improve sexual health outcomes, engagement, and user experience with such an intervention.

Conclusions and Implications for Design

The mobile app of Anathema with a sexual health promotion program delivered in a self-guided mode to a sample of older adults in the Netherlands was assessed as usable. Most participants tended to assess the app and program as useful, but both the app and the program would benefit from certain improvements, which we group under "content" and "interaction" as possible guidelines.

Content wise, readability and engagement can be improved by using plainer language in general, revising sections that sound "too scientific" (eg, definition of sexual health) or too medical (eg, content regarding erectile problems was very focused on the urological aspects). Although for some participants, the content was too long, and participants would also prefer not to have locked content, in which case the length would not be a barrier to engagement. It is clear that participants would appreciate more curated content that would refer them to support the community or to further services.

Regarding *interaction*, there are suggestions to enable searching and asking questions so that the user could be directly guided to the content that is of most interest to them or so that they could center their learning in their own experiences. This implies that content is unlocked by default. Participants would also like to bookmark certain sections and have the means to know where they are in the app. Finally, although participants shared that some exercises made them reflect—they saw this as positive—they struggled with the exercises that involved text input. Therefore, the interaction modes in the exercises can be improved. Although participants state they do not want to be schooled, they highlight "learning" as one of the advantages of using the app. In the future, the Anathema app should meet the goal of teaching without resembling a schoolbook. This was highlighted by participants who were expecting more interactivity from the app rather than an app that reads like a book.

Acknowledgments

Beyond the participants in this study, ≥1400 older adults from 3 different countries were involved in the design and testing of the program. The authors are very thankful to all the participants and to the secondary and tertiary users who provided valuable feedback.

The authors would like to acknowledge the European Commission through the Active Assisted Living (AAL) Programme (reference AAL-2020-7-133-CP) and the Member States: Fundação para a Ciência e a Tecnologia, in Portugal; the Austrian Research Promotion Agency (FFG), in Austria; and ZonMw, the Netherlands Organisation for Health Research and Development, in the Netherlands for funding this project.

Conflicts of Interest

None declared.

XSL•FO
RenderX

Multimedia Appendix 1
Posttest questionnaire.
[DOCX File , 45 KB - humanfactors_v11i1e56206_app1.docx]

Multimedia Appendix 2
Semistructured interview script.
[DOCX File , 39 KB - humanfactors_v11i1e56206_app2.docx]

References

1. World Health Organization. Defining sexual health: report of a technical consultation on sexual health. World Health Organization. 2006. URL: https://tinyurl.com/ye2aep8c [accessed 2024-03-11]
2. Vasconcelos P, Paúl C, Serruya SJ, Ponce de León RG, Nobre P. A systematic review of sexual health and subjective well-being in older age groups. Rev Panam Salud Publica 2022;46:e179 [FREE Full text] [doi: 10.26633/RPSP.2022.179] [Medline: 36320206]
3. Træen B, Hald GM, Graham CA, Enzlin P, Janssen E, Kvalem IL, et al. Sexuality in older adults (65+)—an overview of the literature, part 1: sexual function and its difficulties. Int J Sex Health 2016 Sep 14;29(1):1-10. [doi: 10.1080/19317611.2016.1224286]
4. Towler LB, Graham CA, Bishop F, Hinchliff S. Older adults' embodied experiences of aging and their perceptions of societal stigmas toward sexuality in later life. Soc Sci Med 2021 Oct;287:114355. [doi: 10.1016/j.socscimed.2021.114355] [Medline: 34474307]
5. Sinković M, Towler L. Sexual aging: a systematic review of qualitative research on the sexuality and sexual health of older adults. Qual Health Res 2019 Jul 25;29(9):1239-1254. [doi: 10.1177/1049732318819834] [Medline: 30584788]
6. Wang V, Depp CA, Ceglowski J, Thompson WK, Rock D, Jeste DV. Sexual health and function in later life: a population-based study of 606 older adults with a partner. Am J Geriatr Psychiatry 2015 Mar;23(3):227-233. [doi: 10.1016/j.jagp.2014.03.006] [Medline: 24840610]
7. Bertschi I, Meier F, Bodenmann G. Disability as an interpersonal experience: a systematic review on dyadic challenges and dyadic coping when one partner has a chronic physical or sensory impairment. Front Psychol 2021;12:624609 [FREE Full text] [doi: 10.3389/fpsyg.2021.624609] [Medline: 33732189]
8. Smith L, Yang L, Veronese N, Soysal P, Stubbs B, Jackson SE. Sexual activity is associated with greater enjoyment of life in older adults. Sex Med 2019 Mar;7(1):11-18 [FREE Full text] [doi: 10.1016/j.esxm.2018.11.001] [Medline: 30554952]
9. Landers S, Kapadia F. The public health of pleasure: going beyond disease prevention. Am J Public Health 2020 Feb;110(2):140-141. [doi: 10.2105/ajph.2019.305495]
10. Hunsaker A, Hargittai E. A review of internet use among older adults. New Media Soc 2018 Jul 16;20(10):3937-3954. [doi: 10.1177/1461444818787348]
11. van Lankveld J. Internet-based interventions for women's sexual dysfunction. Curr Sex Health Rep 2016 Jun 30;8(3):136-143 [FREE Full text] [doi: 10.1007/s11930-016-0087-9] [Medline: 27512346]
12. Martin P, Cousin L, Gottot S, Bourmaud A, de La Rochebrochard E, Alberti C. Participatory interventions for sexual health promotion for adolescents and young adults on the internet: systematic review. J Med Internet Res 2020 Jul 31;22(7):e15378 [FREE Full text] [doi: 10.2196/15378] [Medline: 32735217]
13. Karim H, Choobineh H, Kheradbin N, Ravandi MH, Naserpor A, Safdari R. Mobile health applications for improving the sexual health outcomes among adults with chronic diseases: a systematic review. Digit Health 2020 Feb 20;6:2055207620906956 [FREE Full text] [doi: 10.1177/2055207620906956] [Medline: 32128234]
14. Mlekus L, Bentler D, Paruzel A, Kato-Beiderwieden AL, Maier GW. How to raise technology acceptance: user experience characteristics as technology-inherent determinants. Gr Interakt Org 2020 Aug 06;51(3):273-283. [doi: 10.1007/s11612-020-00529-7]
15. Hornbæk K, Hertzum M. Technology acceptance and user experience: a review of the experiential component in HCI. ACM Trans Comput Hum Interact 2017 Oct 06;24(5):1-30. [doi: 10.1145/3127358]
16. Norman D. The Design of Everyday Things. Cambridge, MA: MIT Press; 1988.
17. Interaction Design Foundation. User experience (UX) design. The Interaction Design Foundation. 2023. URL: https://www.interaction-design.org/literature/topics/ux-design?page=7 [accessed 2023-11-27]
18. International Organization for Standardization. ISO 9241-11:2018(en), ergonomics of human-system interaction: part 11: usability: definitions and concepts. International Organization for Standardization. URL: https://www.iso.org/obp/ui/fr/#iso:std:iso:9241:-11:ed-2:v1:en:sec:A [accessed 2023-11-27]
19. van Acker J, Maenhout L, Compernolle S. Older adults' user engagement with mobile health: a systematic review of qualitative and mixed-methods studies. Innov Aging 2023;7(2):igad007 [FREE Full text] [doi: 10.1093/geroni/igad007] [Medline: 37007638]

XSL•FO
RenderX

20. Hurmuz MZ, Jansen-Kosterink SM, Beinema T, Fischer K, Op den Akker H, Hermens HJ. Evaluation of a virtual coaching system eHealth intervention: a mixed methods observational cohort study in the Netherlands. Internet Interv 2022 Mar;27:100501 [FREE Full text] [doi: 10.1016/j.invent.2022.100501] [Medline: 35198411]
21. Abascal J, Nicolle C. Moving towards inclusive design guidelines for socially and ethically aware HCI. Interact Comput 2005 Sep;17(5):484-505. [doi: 10.1016/j.intcom.2005.03.002]
22. Kascak LR, Rébola CB, Sanford JA. Integrating universal design (UD) principles and mobile design guidelines to improve design of mobile health applications for older adults. In: Proceedings of the 2014 IEEE International Conference on Healthcare Informatics. 2014 Presented at: ICHI '14; September 15-17, 2014; Verona, Italy p. 343-348 URL: https://ieeexplore.ieee.org/document/7052509 [doi: 10.1109/ichi.2014.54]
23. Correia de Barros D, Couto da Silva J, Ramadani N, Mendes Santos C. Towards the inclusive design of a digital sexual health promotion programme. In: Proceedings of the 2014 IEEE International Conference on Healthcare Informatics. 2014 Presented at: ICHI '14; September 15-17, 2014; Verona, Italy p. 18 URL: https://tinyurl.com/52tj29jj
24. Mendes-Santos C, Quinta-Gomes AL, Pereira R, Vasconcelos P, Nobre P, Couto J, et al. A smartphone-delivered program (Anathema) to promote the sexual health of older adults, colorectal cancer survivors, and stroke survivors: protocol for a feasibility pilot randomized controlled trial. JMIR Res Protoc 2023 Jun 27;12:e46734 [FREE Full text] [doi: 10.2196/46734] [Medline: 37368469]
25. Bødker S, Kyng M. Participatory design that matters—facing the big issues. ACM Trans Comput Hum Interact 2018 Feb 13;25(1):1-31. [doi: 10.1145/3152421]
26. Mendes SC, Ramadani N, Pereira R, Vasconcelos P, Quinta-Gomes A, Nobre P. Older adults' acceptability of smartphone-based sexual health promotion programs. In: Proceedings of the 11th Scientific Meeting of the International Society for Research on Internet Interventions. 2022 Presented at: ISRII '22; September 18-21, 2022; Pittsburgh, PA URL: https://anathemaproject.eu/wp-content/uploads/2022/10/Anathema_poster_ISRII-scaled.jpg
27. Ribeiro J, Lima P, Nunes F. Trial monitor: scaffolding personalised web dashboards for human–computer interaction field trials. SoftwareX 2021 Dec;16:100883. [doi: 10.1016/j.softx.2021.100883]
28. Ensink CJ, Keijsers NL, Groen BE. Translation and validation of the system usability scale to a Dutch version: D-SUS. Disabil Rehabil 2024 Jan 27;46(2):395-400. [doi: 10.1080/09638288.2022.2160837] [Medline: 36573399]
29. The WHOQOL group. Development of the World Health Organization WHOQOL-BREF quality of life assessment. Psychol Med 1998 May 01;28(3):551-558. [doi: 10.1017/s0033291798006667] [Medline: 9626712]
30. Bangor A, Kortum P, Miller J. Determining what individual SUS scores mean: adding an adjective rating scale. J Usability Stud Exp 2009;4(3):23.
31. Kornfield R, Stamatis CA, Bhattacharjee A, Pang B, Nguyen T, Williams JJ, et al. A text messaging intervention to support the mental health of young adults: user engagement and feedback from a field trial of an intervention prototype. Internet Interv 2023 Dec;34:100667 [FREE Full text] [doi: 10.1016/j.invent.2023.100667] [Medline: 37746639]
32. Ribeiro J, Correia de Barros A. Efficiency of a video and a tutorial in teaching older adults to interact with smartphones. In: Proceedings of the 8th International Conference, UAHCI 2014, Held as Part of HCI International 2014 on Universal Access in Human-Computer Interaction: Aging and Assistive Environments. 2014 Presented at: UAHCI HCI '14; June 22-27, 2014; Crete, Greece p. 34-45. [doi: 10.1007/978-3-319-07446-7_4]
33. Hays RD, Hayashi T, Stewart AL. A five-item measure of socially desirable response set. Educ Psychol Meas 2016 Sep 07;49(3):629-636. [doi: 10.1177/001316448904900315]

Abbreviations

KBO-PCOB: Katholieke Bond van Ouderen - Protestant Christelijke Ouderen Bond
SUS: System Usability Scale

Edited by D Gooch; submitted 11.01.24; peer-reviewed by C Klein, C Abbatantuono, A Batrakoulis, S Markham; comments to author 09.02.24; revised version received 22.02.24; accepted 22.02.24; published 03.04.24.

Please cite as:
Correia de Barros A, Bergmans M, Hasanaj K, Krasniqi D, Nóbrega C, Carvalho Carneiro B, Vasconcelos PA, Quinta-Gomes AL, Nobre PJ, Couto da Silva J, Mendes-Santos C
Evaluating the User Experience of a Smartphone-Delivered Sexual Health Promotion Program for Older Adults in the Netherlands: Single-Arm Pilot Study
JMIR Hum Factors 2024;11:e56206
URL: https://humanfactors.jmir.org/2024/1/e56206
doi:10.2196/56206
PMID:38568726

XSL•FO
RenderX

©Ana Correia de Barros, Mariëtte Bergmans, Kreshnik Hasanaj, Drianë Krasniqi, Catarina Nóbrega, Bruna Carvalho Carneiro, Priscila A Vasconcelos, Ana Luísa Quinta-Gomes, Pedro J Nobre, Joana Couto da Silva, Cristina Mendes-Santos. Originally published in JMIR Human Factors (https://humanfactors.jmir.org), 03.04.2024. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on https://humanfactors.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Designing for Improved Patient Experiences in Home Dialysis: Usability and User Experience Findings From User-Based Evaluation Study With Patients With Chronic Conditions

Anna Aspelund[1], MSc; Paula Valkonen[1], MA; Johanna Viitanen[1], PhD; Virpi Rauta[2,3], PhD

[1]Department of Computer Science, Aalto University, Espoo, Finland

[2]Department of Nephrology, Helsinki University Central Hospital, University of Helsinki, Helsinki, Finland

[3]Abdominal Center, Nephrology, Helsinki University Hospital, Helsinki, Finland

Corresponding Author:
Anna Aspelund, MSc
Department of Computer Science
Aalto University
PO Box 15400
Espoo, FI-00076
Finland
Phone: 358 504763381
Email: anna.aspelund@aalto.fi

Abstract

Background: Chronic kidney disease affects 10% of the population worldwide, and the number of patients receiving treatment for end-stage kidney disease is forecasted to increase. Therefore, there is a pressing need for innovative digital solutions that increase the efficiency of care and improve patients' quality of life. The aim of the eHealth in Home Dialysis project is to create a novel eHealth solution, called eC4Me, to facilitate predialysis and home dialysis care for patients with chronic kidney disease.

Objective: Our study aimed to evaluate the usability, user experience (UX), and patient experience (PX) of the first version of the eC4Me solution.

Methods: We used a user-based evaluation approach involving usability testing, questionnaire, and interview methods. The test sessions were conducted remotely with 10 patients with chronic kidney disease, 5 of whom had used the solution in their home environment before the tests, while the rest were using it for the first time. Thematic analysis was used to analyze user test and questionnaire data, and descriptive statistics were calculated for the UMUX (Usability Metric for User Experience) scores.

Results: Most usability problems were related to navigation, the use of terminology, and the presentation of health-related data. Despite usability challenges, UMUX ratings of the solution were positive overall. The results showed noteworthy variation in the expected benefits and perceived effort of using the solution. From a PX perspective, it is important that the solution supports patients' own health-related goals and fits with the needs of their everyday lives with the disease.

Conclusions: A user-based evaluation is a useful and necessary part of the eHealth solution development process. Our study findings can be used to improve the usability and UX of the evaluated eC4Me solution. Patients should be actively involved in the solution development process when specifying what information is relevant for them. Traditional usability tests complemented with questionnaire and interview methods can serve as a meaningful methodological approach for gaining insight not only into usability but also into UX- and PX-related aspects of digital health solutions.

(JMIR Hum Factors 2024;11:e53691) doi:10.2196/53691

KEYWORDS

usability; UX; user experience; PX; patient experience; user-based evaluation; patients; eHealth; digital health solution; kidney disease; home dialysis

XSL•FO
RenderX

Introduction

Chronic kidney disease is a global health problem that leads to kidney failure, cardiovascular disease, and premature death. Chronic kidney disease affects 10% of the population and is one of the leading causes of mortality worldwide [1]. Dialysis, along with kidney transplant, is a lifesaving treatment for people with end-stage kidney disease. Dialysis can be delivered in hospital or home settings, and home dialysis is associated with a higher or equal quality of life for patients [2] and lower costs for the health care system [3]. As the number of patients receiving treatment for end-stage kidney disease is forecasted to rise [4], innovative digital solutions that maximize efficiency, improve patients' quality of life, and facilitate care delivery and monitoring are needed.

eHealth solutions, such as digital patient engagement platforms (DPEPs), are increasingly developed to support self-care, enhance patient-clinician collaboration, and increase the efficiency of care delivery [5-7]. In dialysis care, new DPEP solutions have the potential to improve disease management, health outcomes, and patient experience (PX) among patients with chronic conditions [8,9]. To achieve these goals, a human-centered design approach to development is a necessity. Human-centered design is an approach that aims to make digital systems usable and useful by applying human factors and usability techniques, such as user-based testing, guidelines for interaction design, prototypes, user observations, and user requirements specifications [10]. Usability refers to the interaction between the end user and the system, whereas the user experience (UX) includes aspects like emotions, beliefs, and perceptions [10,11]. Originating from the UX, the PX has also become an important and acknowledged concept as the health care sector has shifted to a more customer-oriented approach. PX has been used to describe patients' interactions and care experiences across the care continuum [12,13], but it lacks a consensus definition [14]. Regarding eHealth solutions, numerous factors influence PX, such as the solution type and quality, risks and concerns, communication, remote interaction, and patients' attitudes toward digital solutions [14].

Several studies have evaluated the usability of eHealth solutions aimed at patients with chronic and serious conditions. These have included solutions targeted to patients with cancer for monitoring and managing their illness or treatment-related symptoms [15-17], digital self-management programs for patients with juvenile idiopathic arthritis [18] and chronic obstructive pulmonary disease [19], and an electronic patient-reported outcome tool for patients with complex chronic disease and disability to set and monitor their health-related goals [20]. Common usability problems identified across these studies have included terminology issues [15,18], navigation problems [15,17], and challenges with the way information is presented to the patients [16,18]. Regarding UX, studies have found that patients' illness-related problems and limitations should be taken into account when designing eHealth solutions for patients with chronic and serious conditions [16,19,20]. Further, customization of the solutions, for example, based on the stage or severity of the illness or type of treatment should be possible to provide a pleasant UX [16,17,19]. Some prior studies have also reported PX-related findings, such as patients fearing that the eHealth solutions will replace in-person consultations with clinicians [20], and patients generally welcoming the additional digital communication channel [16,17]. However, these results have not been analyzed or described in relation to PX, and it seems that PX-related aspects were not systematically explored in the evaluation studies.

In this paper, we report a user-based evaluation study of the novel eHealth solution: a DPEP targeted to patients with chronic kidney disease in CKD stages 4-5, for example, to patients undergoing predialysis and patients undergoing home dialysis (both peritoneal dialysis and home hemodialysis). Patients with functioning renal transplants were excluded. This study is part of the larger eHealth in Home Dialysis project [21], which is coordinated by HUS Helsinki University Hospital, Finland. The solution is designed to facilitate advanced home care: enable patients with chronic kidney disease to document their treatment data, monitor their clinical and health data, order dialysis supplies, and report their symptoms as well as enhance patient-provider communication. The objective of our study was to evaluate the first version of the DPEP solution, called eC4Me, and support deployment of the solution and promote end user participation in later phases of the development. The research questions are as follows: (1) What kind of usability problems does the evaluated DPEP solution have? (2) What kind of UXs, expectations, and improvement ideas do patients with chronic kidney disease have about the new DPEP solution? and (3) How can the new DPEP solution support positive PX for patients with chronic conditions?

Our user-based evaluation study aims to widen the scope of usability evaluations of eHealth solutions targeted at patients with chronic and serious conditions to include PX-related aspects alongside usability and UX. Additionally, to our best knowledge, this is the first study to evaluate a DPEP solution specifically targeted to patients with chronic kidney disease.

Methods

Study Design

Our study design was based on a formative user-based evaluation approach [22]. The formative evaluation aims to support the improvements of the system, particularly the user interface, as part of an iterative design process [22]. A typical method of formative evaluation is a think-aloud usability test, which includes 4 stages: preparation, introduction, the test itself, and debriefing [22]. Usability testing is stated to be the most fundamental usability method since it provides direct information about how people use the systems and what their exact problems are [22]. In practice, the think-aloud method involves the test participants continuously thinking aloud while performing the predefined test tasks. The researcher's role is to make observations and continuously prompt the participant to think aloud by asking general questions [22].

The usability assessment methods recommended for gathering supplementary data are observations, questionnaires, and interviews [22]. For the questionnaire, we used the UMUX (Usability Metric for User Experience) [23], which closely

XSL•FO
RenderX

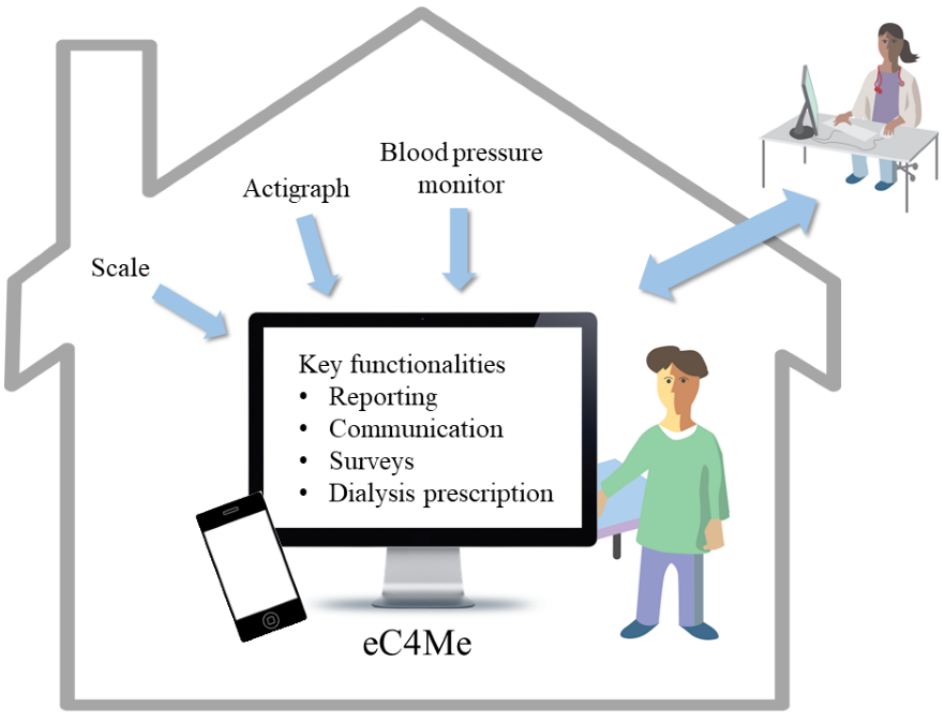
conforms to the 3 widely acknowledged attributes of usability—effectiveness, efficiency, and satisfaction [10]—and strongly correlates with other commonly used usability metrics such as the System Usability Scale [24,25]. UMUX questionnaire is considered compact since it includes four question items: (1) the system's capabilities meet my requirements, (2) using the system is a frustrating experience, (3) the system is easy to use, and (4) I have to spend too much time correcting things with this system [23].

Due to the COVID-19 pandemic, we conducted the evaluation sessions remotely. Experiences from previous studies have shown that high-quality research data can be collected remotely [26]. However, compared to traditional face-to-face usability testing, remote evaluations require more pedantic preparation. Researchers must pay attention to building trust and confidentiality [26], choose tools that are familiar and easily accessible for the participants [27], focus on body language and facial expressions to establish rapport [26,28], and provide the participants with technical support as needed [29]. Other recommendations include having multiple researchers participate in remote test sessions [28], and using a synchronous approach, which makes it possible to observe participants' screens in real-time [30].

Evaluated DPEP Solution

The first version of the eC4Me solution was introduced in autumn 2021. A core part of the solution was an app, which had both computer and mobile interfaces and included the following key functionalities: monitoring of treatment-related data (reporting functionality), messaging between patients and nurses (communication functionality), answering quality-of-life surveys (survey functionality), and access to patient's dialysis prescriptions (dialysis prescription functionality; see Figure 1). In addition, the solution delivered to the patients included external monitoring devices, such as a blood pressure monitor, a scale, and an actigraph, which, together with clinical data collected from electronic health records, enabled patients to monitor their conditions.

Figure 1. First version of the eC4Me solution.



Participants

The participants were recruited from a large university-affiliated nephrology clinic with the help of research nurses. All participants were familiar with the eHealth in Home Dialysis project since they had participated in an interview study that was conducted earlier as part of the project.

Eleven patients with chronic kidney disease were originally invited to participate in this study. As 1 test participant was particularly interested in technology development and came to his test session with predrafted design ideas, we decided to modify his test session to focus on discussing these ideas. Consequently, he did not perform the test tasks in our test procedure, and the data from his session was omitted from this study. Therefore, data from 10 user-based evaluation sessions were included in this study.

The background information collected from the participants included age, gender, type of treatment, occupational status, and technical skills. Technical skills were evaluated by asking participants to give their own estimation of their technical skills, with response options being "good," "basic," and "weak." Half of the participants (n=5) had the solution delivered to their homes 1-3 weeks before the tests, whereas the other half (n=5) were using the solution for the first time in their test session. Participant characteristics are shown in Table 1.

XSL•FO
RenderX

Table 1. Participant characteristics (N=10).

Characteristics	Participants, n (%)
Type of treatment	
Predialysis (not yet in dialysis treatment)	4 (40)
Peritoneal home dialysis	3 (30)
Home hemodialysis	3 (30)
Experience with the solution before the test session	
No experience	5 (50)
1-3 weeks of experience	5 (50)
Sex	
Male	6 (60)
Female	4 (40)
Age (y)	
30-60	4 (40)
>60	6 (60)
Occupational status	
Fully working	2 (20)
Partially working	2 (20)
Not working (fully retired or sick leave)	6 (60)
Technical skills (own estimation)	
Good	2 (20)
Basic	7 (70)
Weak	1 (10)

Test Procedure

In our study, we used synchronous remote usability testing with Microsoft Teams as a tool, and 2 researchers were present in each session. During the test sessions, patients used the eC4Me solution with a computer, enabling the researchers to monitor their task performance via screen share on Microsoft Teams.

Each test session followed the traditional structure and stages of usability testing [22], lasted about 2 hours, and included the following phases: (1) introduction—participants were introduced to the evaluation study and given an opportunity to become familiar with Microsoft Teams; (2) test tasks—run-through of predefined usability test tasks, which included logging in, searching for information and functionalities, viewing and interpreting health-related data, reporting treatment-related data, and filling in the surveys; (3) questionnaire—participants answered the UMUX questionnaire; and (4) interview—participants answered semistructured interview questions to elaborate their UMUX scores and give overall feedback on the solution based on the usability test tasks. The interview consisted of 4 open-ended “Why did you give this score?” questions, which were asked for each of the UMUX items separately, and a question on how participants would improve the solution.

Before the actual tests, the test procedure was piloted with 2 research nurses. To ensure privacy, all patients used the solution with test login IDs and dummy health data during the test sessions. The exact test tasks varied slightly between the participants, depending on their prior experience with the solution, illness stage, and type of treatment, as not all functionalities of the solution were relevant for all patients. The participants who had used the solution before the tests were also encouraged to provide feedback on the entire solution including the research devices and a mobile interface. This study was performed in the Finnish language.

Data Analysis

The qualitative data included video recordings from remote usability tests, observation notes, and transcripts from semistructured interviews. The qualitative data were analyzed following a thematic analysis method [31], which involved collaboration between 3 researchers (AA, PV, and JV). The data were first coded by 1 researcher (PV), and the findings were discussed by the 3 researchers. Further, 2 researchers (AA and JV) then continued the analysis by categorizing the codes into thematic groups, following the principles of the affinity diagram method [32]. The following main three thematic groups were used: (1) *usability*, which includes findings about users’ interactions with the DPEP solution; (2) *UX*, which includes findings about users’ experiences and feelings toward using the DPEP solution; and (3) *PX*, which includes findings about how the DPEP solution can support patients’ interactions and care experiences across the care continuum.

XSL•FO
RenderX

The researchers (AA and JV) then continued the analysis with several rounds of iterations. Along with other data, improvement ideas expressed by the participants were thematically grouped. At the end of the analysis, the thematic grouping of observations was discussed, approved, and finalized collaboratively by the 3 researchers.

The quantitative data consisted of UMUX item scores, which were analyzed following the UMUX scoring scheme [23]: to obtain the overall UMUX score, items 1 and 3 were scored as [score−1] and items 2 and 4 as [7−score], and the sum of the item scores was then divided by 24 and multiplied by 100. In addition to the overall score, the means and SDs for each of 4 question items were calculated separately for 2 participant groups (patients who had or had not used the solution before the test). The differences between the groups were analyzed using *t* tests for independent samples. The tools used for data analysis were ATLAS.ti (ATLAS.ti Scientific Software Development GmbH) and Microsoft Excel for qualitative data analysis and Microsoft Excel for statistical analysis.

Ethical Considerations

This study has a research permit from the ethical committee of the Hospital District of Helsinki and Uusimaa (HUS/1649/2020).

Results

Overview

The results are divided into 5 topics: usability, UX and PX findings each, UMUX results, and improvement ideas.

Usability

The *usability* findings of the evaluated eC4Me solution consisted of 8 subthemes (Table 2).

Navigation includes findings about whether patients could locate the functionalities, content, and commands that they were looking for. Nine out of 10 users had at least some problems navigating the app, and the most common navigation challenges were related to users not understanding the content structure or the terminology used in the menus.

Table 2. *Usability*, *UX*, and *PX* findings of the user-based evaluation. "All findings" includes positive, negative, and neutral findings. For the *usability* theme, negative findings, that is, the identified usability problems, are also reported separately under "problems."

Subtheme	All findings		Problems	
	Codes, n	Users, n	Codes, n	Users, n
Usability				
Navigation	41	10	19	9
Terminology	35	10	22	8
Front page	21	10	3	2
Presentation of data	35	9	17	5
Login	10	9	4	4
Survey functionality[a]	12	7	2	2
Reporting functionality[a]	50	6	29	6
Dialysis prescription functionality[a]	6	5	0	0
UX[b]				
Technical functionality	26	9	—[c]	—
Use of access devices (computer, tablet, or mobile)	26	9	—	—
Workload and effort	16	7	—	—
Perceived benefits	12	7	—	—
Security	8	4	—	—
PX[d]				
Content-related needs	67	10	—	—
Situation of use	57	10	—	—
Communication with clinicians	52	10	—	—

[a]The survey, reporting, and dialysis prescription functionalities were tested with some of the participants only (n=7, n=6, and n=5, respectively).

[b]UX: user experience

[c]Not applicable.

[d]PX: patient experience

XSL•FO
RenderX

Terminology includes findings about the comprehensibility and clarity of the terminology used. Eight users (80%) had problems understanding the terminology, and approximately half (10/22) of the terminology challenges were related to problems with understanding medical- or treatment-related terminology. Other terminology issues included problems with the terms used in the menus as well as the use of a foreign language.

Front page includes findings about the comprehensibility and clarity of the front-page contents. The front page of the tested version contained relatively little information and functionalities, and most users found it simple and clear.

Presentation of data includes findings about the comprehensibility and clarity of the presentation of health data, such as health measurements. Five users (50%) had issues understanding or viewing the data. The most common challenges were not comprehending the data or graphs or not knowing how to adjust the scales and timelines to view the data in a meaningful way.

Login includes findings about the ease of logging in. Four users (40%) had problems logging into the system. Typical challenges included not understanding where to input the login information or making errors while typing the login details.

Surveys, reporting, and dialysis prescription functionalities include findings about the ease of use of these functionalities. All users who tested the reporting functionality (6/6) had problems using it. Users struggled with not understanding what they should type in the input fields, feeling that options in the fields did not match the way treatment was provided in the real world, or not comprehending the medical- or treatment-related terminology. In this study, there were few usability issues in the survey functionality and none in the dialysis prescription functionality.

About UX

The *UX* findings of the evaluated eC4Me solution consisted of 5 subthemes (Table 2).

Technical functionality includes patients' experiences and feedback regarding the technical aspects of the eC4Me solution. Four users (40%) expressed frustration because some information they thought should be transferred automatically between the app, the research devices, the home dialysis machine, and patient information systems had to be typed manually. For the same reason, 2 users (20%) felt that they needed to use several systems for essentially the same purpose, such as monitoring their health data.

Use of access devices includes patients' expressed preferences regarding using the solution with different access devices: desktop computer, tablet, or mobile phone. Two users (20%) said they would prefer to use the computer interface, as they have found tablet and mobile keyboards difficult to use or feel that the mobile interface would give them less information. In contrast, 3 (30%) users indicated that they preferred a mobile phone or tablet as they are readily at hand and easier to use during the treatment, while another 3 (30%) said that their choice of access device would depend on the task they were performing.

Workload and effort includes findings about the perceived time and effort required to use the solution. Six users (60%) felt that the solution was not burdensome to use as such and that filling in the surveys or documenting treatment details did not take too much time. However, 4 users (40%) expressed concern that the solution might nevertheless increase their burden if it does not replace any other service, thus becoming one more thing to use and keep track of on top of all the other health-related solutions.

Perceived benefits includes patients' thoughts about the benefits and added value of the eHealth solution. One (10%) user saw value in using the solution primarily for the benefit of the health care personnel, while 2 (20%) others said that they needed to see clear benefits for themselves to be motivated to use the solution. Yet another user mentioned that the data generated by the solution could benefit all patients, as it could be used for research and treatment development.

> *I'm uncertain what this is meant for, is it for my benefit or someone else's? The remote measuring devices that I have had, I have found the data very useful for myself.... But I don't understand the thinking behind this (the solution), do I benefit or is it someone else?* [P7]

Security includes findings about potential security issues and patients' concerns regarding the use of the eHealth solution. Only 2 (20%) users gave direct comments on security aspects, while most findings related to security were observations of behaviors that could introduce potential security risks, such as the user closing the browser instead of logging out when asked to do so.

About PX

The *PX* findings of the evaluated eC4Me solution consisted of 3 subthemes (Table 2).

Content-related needs includes patients' comments regarding health-related data that they want to see so they can monitor and manage their treatment and health. The expressed needs and what was considered most important varied between the users, but overall, patients were interested in seeing all the types of data that the tested version of the solution provided. Only 1 user (10%) gave a general comment that the solution "should not contain anything unnecessary or useless," but other than that, none of the users reported that they would not need or want to see some of the information or data that was available to them.

Situation of use includes patients' comments and feedback about how well the solution fits their situations and supports their everyday lives with the disease. Users had numerous, often variable comments regarding how often and in what situations they would likely use the solution. They also commented on how well the functionalities fit their care and treatment schedules, as in the following quote:

> *I fill these during my home dialysis treatment, so I may write notes about yesterday's treatment. I don't necessarily have time to use [the solution] after the treatment.* [P8]

XSL•FO
RenderX

Communication with clinicians includes findings about how the new solution supports patient-clinician communication. Users expressed interest in using the messaging function and saw benefits in using the documented data to facilitate their communication with clinicians during face-to-face appointments. It was not clear to the users how actively and by whom their data were being monitored and if messages were noticed and replied to. Three users (30%) were hoping for immediate feedback, while 4 others (40%) considered the messaging function appropriate for nonurgent communication. In addition, 5 users (50%) expected their own nurse to read and respond to their messages, while 3 users (30%) thought that the work was handled by a care team.

UMUX Results

The UMUX score of the first version of eC4Me was 70.6 (SD 18.6), which indicates an average level of usability [25].

Means for individual UMUX items are presented in Table 3. Users with 1-3 weeks of prior experience with the solution rated it more favorable overall compared to users without prior experience. However, the differences between the groups were not statistically significant.

Table 3. UMUX[a] item scores per user groups on a scale of 1 "strongly disagree" to 7 "strongly agree."

UMUX questionnaire item	All users (n=10), mean (SD) score	Users with no prior experience using the solution (n=5), mean (SD) score	Users with 1-3 weeks of experience using the solution (n=5), mean (SD) score
The solution's capabilities meet my requirements	4.5 (1.4)	4.0 (1.6)	5.0 (1.0)
Using the solution is a frustrating experience	2.5 (1.9)	3.0 (2.3)	1.9 (3.0)
The solution is easy to use	5.7 (1.5)	5.0 (1.9)	6.4 (0.5)
I have to spend too much time correcting things with the solution	2.8 (1.9)	3.0 (1.4)	2.6 (2.5)

[a]UMUX: Usability Metric for User Experience.

Improvement Ideas

In total, 66 improvement ideas (Table 4) for the eC4Me solution were identified from the data, with all 10 users expressing at least one improvement idea. Two-thirds of the ideas (40/66) came from users who had used the solution before the test.

The most common theme for improvement ideas was *content-related needs*. Seven patients (70%) expressed interest in monitoring some health-related measurements that were not available in the tested version, and 3 (30%) patients wanted to see benchmark values or descriptions that would enable them to better understand their health data.

Table 4. Improvement ideas and their most common subthemes.

Subtheme	All improvement ideas	
	Codes, n	Users, n
Content-related needs	28	9
Situation of use	15	7
Communication with clinicians	15	6
Presentation of data	8	6
Ease of using reporting	13	5
Technical functionality	8	5
Other[a]	26	8

[a]Subthemes with fewer than 5 ideas (combined).

In addition, the participants brought up improvement ideas related to the following: (1) *situation of use*—ideas on how the solution could be improved to better fit the patient's situation, everyday life, and treatment schedule; (2) *communication with clinicians*—ideas on how the solution could better support communication and data exchange between patients and clinicians; (3) *presentation of data*—ideas on how health data could be presented to make them more meaningful for the patients; (4) *ease of using reporting*—ideas on how to improve the reporting functionality to make it easier to use; and (5) *technical functionality*—ideas regarding automatic data exchange between the solution and other devices or services.

Discussion

Main Contribution

Our user-based evaluation study of the novel DPEP solution targeted to patients with chronic kidney disease with 10 participants resulted in a wide variety of usability-, UX-, and PX-related findings.

Most usability problems of the first version of the solution were related to navigation, the use of terminology, and the presentation of health data. Many participants struggled with the reporting functionality, which was one of the key functionalities of the solution. A considerable number of patient

XSL•FO
RenderX

participants also expressed improvement ideas related to these themes. We decided not to classify usability problems by severity, as a proper severity rating should consider not only usability aspects but also potential medical- and health-related consequences of users' mistakes and misunderstandings. However, the usability challenges identified in our study were remarkably similar to those found in evaluations of other eHealth solutions aimed at patients with chronic and serious conditions [15-18]. Our findings thus emphasize the importance of using terminology and presenting health data in a way that is understandable and meaningful to patients. Our results also highlight the need to consider patients as end users when designing user interfaces for eHealth solutions.

Our study identified several challenges related to the UX of the evaluated DPEP solution. Largely due to deficiencies in integration and data exchange, the participants feared that the solution might create additional tasks and thus increase their burden. Our results also showed considerable variation in the expected benefits of the solution. Some patients wanted to see direct value for themselves, whereas others mentioned benefits for the health care professionals as their primary motivation of use.

Despite the usability and UX challenges, the patients' overall ratings of the evaluated solution were surprisingly positive. This may be at least partially explained by findings from previous studies, which have shown that patients with chronic and serious conditions often express high interest in disease-specific eHealth solutions [17,18], even when experiencing severe usability challenges [20]. In our study, patients who had used the solution for a few weeks in a home setting evaluated it more positively than patients who were using the solution for the first time. Although the differences were not statistically significant due to the small number of participants, these initial findings could simply be explained by the fact that learning to use the solution makes it easier and thus more pleasant to use. However, they could also indicate that after having used the solution in their home setting with their own health data, patients have a better understanding of the benefits and potential value of the solution.

Regarding PX, our study generated insights on how the DPEP solution can support patients with chronic conditions in monitoring and managing their conditions and how the DPEP solution could better fit their everyday lives with a disease. From the patients' perspective, it is not enough that an eHealth solution is easy or pleasant to use if it does not support their health-related goals, feel meaningful, and fit their real-life situations and daily care activities. Special attention needs to be paid to ensure that these PX-related considerations are included in user-based evaluations of eHealth solutions, as generic usability questionnaires, classifications, and frameworks do not adequately capture these aspects [33,34]. As our study shows, traditional usability tests, complemented with questionnaire and related interview methods, can serve as a meaningful methodological approach for collecting information about PX-related aspects of eHealth solutions.

In our study, the participants generated a considerable amount of improvement ideas. In particular, nearly all patients had ideas on what health-related data they would like to see to better manage their condition. This implies 2 things. First, many patients with chronic conditions are interested in taking responsibility for their own care. Although the participants selected for our study are likely to represent the most motivated and active patients with chronic kidney disease, it would seem meaningful to support and empower these motivated patients to take more responsibility by providing them with the information they view as important and meaningful, not only the information that makes the most sense from a health care professionals' point of view. Second, patients should be actively involved in the co-design process in the early phases of solution development and when specifying what kind of information is relevant for them.

The user-based evaluation was a crucial step in the eHealth solution development process and generated findings that helped to make substantial changes in the solution to make it more suitable for the end users (patients), thus helping the solution reach its goals. The evaluation of usability-, UX-, and PX-related aspects of the solution will continue in a future research project. We aim to conduct a similar study in a further phase of the development project to examine how the usability and UX of the solution have been improved.

Limitations

Due to the COVID-19 pandemic, our study used remote testing as an evaluation method. In comparison to in-context evaluations, this limited the scope of our evaluation, as we could not fully observe participants in their home environments. We also decided not to include research devices that were part of the DPEP solution in the test procedure, as this would have been difficult to realize in the remote setup. However, when compared to face-to-face usability testing in laboratory settings, our arrangement also had some advantages. As contextual factors are well known to influence emotional experiences and expressions [35], allowing patients to remain in their natural home settings during test sessions likely produced more reliable data, especially regarding experience-related topics like UX and PX.

Further Research

Our findings indicate that including participants who have used the evaluated solution before the test can have a nonnegligible effect on the quality and amount of information that the evaluation study generates. Half of the participants in our study had used the evaluated solution in their home environment with their own health data, while the other half were using the solution for the first time in their test session. As the number of participants was small (n=10) and the groups were heterogeneous in terms of other background variables, it was not meaningful to make more comprehensive comparisons between the 2 groups. However, participants with prior experience evaluated the solution more favorably and generated more improvement ideas. Many UX- and PX-related aspects, such as perceived benefits, workload, and compatibility with everyday life, can be difficult to assess using a solution only in a test setting, especially considering that privacy issues often prevent researchers from using patients' own health data in user tests. This could have implications on how user-based evaluation

XSL•FO
RenderX

studies of eHealth solutions should ideally be arranged, and it is therefore an important topic for further research.

In addition, further research is needed to explore the relationship and connections between the concepts of UX and PX, as suggested by recent review studies [14,36]. This includes planning and practicalities of user-based evaluation studies, considering the PX perspective, for the assessment and improvement of eHealth services.

Conclusions

User-based evaluation can produce valuable findings about usability aspects but also about the UX and PX of the evaluated DPEP solution. The findings of our study can be used in the development process to improve the evaluated solution from the perspective of patients with chronic conditions. Evaluation is a useful and necessary part of the solution development process, especially considering the high number of novel eHealth solutions that are currently being developed.

Our study also highlights the importance of understanding how digital health solutions for patients with chronic and serious conditions support patients' own health-related goals and fit their lives with disease. To fully understand the motivation for using such solutions, it is necessary to understand how patients perceive the benefits versus the effort required to use the solution in their everyday lives.

Acknowledgments

The eHealth in Home Dialysis project was part of the CleverHealth Network ecosystem. The project was funded by participating companies, the Hospital District of Helsinki and Uusimaa (HUS), and Business Finland. We thank our participants, project partners, and colleagues: S Hölsä, S Leinonen, and N Karisalmi from Aalto University and T Puurtinen, P Silvast-Äikäs, and A Varila from HUS.

Conflicts of Interest

None declared.

References

1. Kovesdy CP. Epidemiology of chronic kidney disease: an update 2022. Kidney Int Suppl 2022;12(1):7-11 [FREE Full text] [doi: 10.1016/j.kisu.2021.11.003] [Medline: 35529086]
2. Bonenkamp AA, van Eck van der Sluijs A, Hoekstra T, Verhaar MC, van Ittersum FJ, Abrahams AC, et al. Health-related quality of life in home dialysis patients compared to in-center hemodialysis patients: a systematic review and meta-analysis. Kidney Med 2020;2(2):139-154 [FREE Full text] [doi: 10.1016/j.xkme.2019.11.005] [Medline: 32734235]
3. Howell M, Walker RC, Howard K. Cost effectiveness of dialysis modalities: a systematic review of economic evaluations. Appl Health Econ Health Policy 2019;17(3):315-330. [doi: 10.1007/s40258-018-00455-2] [Medline: 30714086]
4. Liyanage T, Ninomiya T, Jha V, Neal B, Patrice HM, Okpechi I, et al. Worldwide access to treatment for end-stage kidney disease: a systematic review. Lancet 2015;385(9981):1975-1982. [doi: 10.1016/S0140-6736(14)61601-9] [Medline: 25777665]
5. Shaw T, McGregor D, Brunner M, Keep M, Janssen A, Barnet S. What is eHealth (6)? Development of a conceptual model for eHealth: qualitative study with key informants. J Med Internet Res 2017;19(10):e324 [FREE Full text] [doi: 10.2196/jmir.8106] [Medline: 29066429]
6. Campbell K, Louie P, Levine B, Gililland J. Using patient engagement platforms in the postoperative management of patients. Curr Rev Musculoskelet Med 2020;13(4):479-484 [FREE Full text] [doi: 10.1007/s12178-020-09638-8] [Medline: 32388724]
7. Hölsä S, Viitanen J, Valkonen P, Lääveri T, Rauta V. Developing eHealth for Home Dialysis: Clinicians' Needs for a Digital Patient Engagement Platform. Studies in health technology and informatics 2024;310:1111-1115. [doi: 10.3233/shti231137]
8. Shen H, van der Kleij RMJJ, van der Boog PJM, Chang X, Chavannes NH. Electronic health self-management interventions for patients with chronic kidney disease: systematic review of quantitative and qualitative evidence. J Med Internet Res 2019;21(11):e12384 [FREE Full text] [doi: 10.2196/12384] [Medline: 31687937]
9. Cartwright EJ, Zs Goh Z, Foo M, Chan CM, Htay H, Griva K. eHealth interventions to support patients in delivering and managing peritoneal dialysis at home: a systematic review. Perit Dial Int 2021;41(1):32-41 [FREE Full text] [doi: 10.1177/0896860820918135] [Medline: 32301380]
10. ISO 9241-210:2019 ergonomics of human-system interaction—part 210: human-centred design for interactive systems. International Organization for Standardization. URL: https://www.iso.org/standard/77520.html [accessed 2023-07-19]
11. Hassenzahl M. User experience and experience design. In: Soegaard M, Dam RF, editors. The Encyclopedia of Human-Computer Interaction, 2nd Edition. Aarhus, Denmark: The Interaction Design Foundation Publisher; 2013.
12. Wolf JA, Niederhauser V, Marshburn D, LaVela SL. Defining patient experience. Patient Exp J 2014;1(1):7-19. [doi: 10.1097/naq.0000000000000293]
13. Meyer MA. COVID-19 pandemic accelerates need to improve online patient engagement practices to enhance patient experience. J Patient Exp 2020;7(5):657-664 [FREE Full text] [doi: 10.1177/2374373520959486] [Medline: 33294595]

XSL•FO
RenderX

14. Viitanen J, Valkonen P, Savolainen K, Karisalmi N, Hölsä S, Kujala S. Patient experience from an eHealth perspective: a scoping review of approaches and recent trends. Yearb Med Inform 2022;31(1):136-145 [FREE Full text] [doi: 10.1055/s-0042-1742515] [Medline: 36463871]
15. Hanghøj S, Boisen KA, Hjerming M, Elsbernd A, Pappot H. Usability of a mobile phone app aimed at adolescents and young adults during and after cancer treatment: qualitative study. JMIR Cancer 2020;6(1):e15008 [FREE Full text] [doi: 10.2196/15008] [Medline: 31895046]
16. Mirkovic J, Kaufman DR, Ruland CM. Supporting cancer patients in illness management: usability evaluation of a mobile app. JMIR mHealth uHealth 2014;2(3):e33 [FREE Full text] [doi: 10.2196/mhealth.3359] [Medline: 25119490]
17. Moradian S, Krzyzanowska MK, Maguire R, Morita PP, Kukreti V, Avery J, et al. Usability evaluation of a mobile phone-based system for remote monitoring and management of chemotherapy-related side effects in cancer patients: mixed-methods study. JMIR Cancer 2018;4(2):e10932 [FREE Full text] [doi: 10.2196/10932] [Medline: 30578238]
18. Stinson J, McGrath P, Hodnett E, Feldman B, Duffy C, Huber A, et al. Usability testing of an online self-management program for adolescents with juvenile idiopathic arthritis. J Med Internet Res 2010;12(3):e30 [FREE Full text] [doi: 10.2196/jmir.1349] [Medline: 20675293]
19. Voncken-Brewster V, Moser A, van der Weijden T, Nagykaldi Z, de Vries H, Tange H. Usability evaluation of an online, tailored self-management intervention for chronic obstructive pulmonary disease patients incorporating behavior change techniques. JMIR Res Protoc 2013;2(1):e3 [FREE Full text] [doi: 10.2196/resprot.2246] [Medline: 23612363]
20. Gray CS, Gill A, Khan AI, Hans PK, Kuluski K, Cott C. The electronic patient reported outcome tool: testing usability and feasibility of a mobile app and portal to support care for patients with complex chronic disease and disability in primary care settings. JMIR mHealth uHealth 2016;4(2):e58 [FREE Full text] [doi: 10.2196/mhealth.5331] [Medline: 27256035]
21. Home dialysis. CleverHealth Network. URL: https://www.cleverhealth.fi/en/development-of-home-dialysis/ [accessed 2023-07-19]
22. Nielsen J. Usability Engineering. San Francisco, CA: Morgan Kaufmann; 1994.
23. Finstad K. The usability metric for user experience. Interact Comput 2010;22(5):323-327. [doi: 10.1016/j.intcom.2010.04.004]
24. Borsci S, Federici S, Bacci S, Gnaldi M, Bartolucci F. Assessing user satisfaction in the era of user experience: comparison of the SUS, UMUX, and UMUX-LITE as a function of product experience. Int J Hum-Comput Interact 2015;31(8):484-495. [doi: 10.1080/10447318.2015.1064648]
25. Lewis JR. Measuring perceived usability: the CSUQ, SUS, and UMUX. Int J Human-Computer Interact 2018;34(12):1148-1156. [doi: 10.1080/10447318.2017.1418805]
26. Valkonen P, Karisalmi N, Kaipio J, Kujala S. Remote interviews and visual timelines with patients: lessons learned. Stud Health Technol Inf 2021;281:845-849. [doi: 10.3233/shti281]
27. Amberkar S, Bennett K, Delchamps A, Edwards R. Expanding the playtest method for UX research beyond gaming. Proc Hum Factors Ergon Soc Annu Meet 2019;63(1):2239-2243. [doi: 10.1177/1071181319631119]
28. Judge TK, Neustaedter C. Studying and Designing Technology for Domestic Life: Lessons from Home, 1st Edition. Burlington, MA: Morgan Kaufmann; 2014.
29. Petrie H, Hamilton F, King N, Pavan P. Remote usability evaluations with disabled people. 2006 Presented at: CHI '06: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems; April 22-27, 2006; Montreal, QC p. 1133-1141 URL: https://dl.acm.org/doi/proceedings/10.1145/1124772 [doi: 10.1145/1124772.1124942]
30. Vasalou A, Ng B, Wiemer-Hastings P, Oshlyansky L. Human-moderated remote user testing: protocols and applications. In: Proceedings of the 8th ERCIM Workshop, User Interfaces for All. 2004 Presented at: 8th ERCIM Workshop, User Interfaces for All; June 28-29, 2004; Vienna, Austria URL: https://reed.cs.depaul.edu/peterh/papers/b2hd-VasalouERCIM2004.html
31. Braun V, Clarke V. Using thematic analysis in psychology. Qual Res Psychol 2006;3(2):77-101. [doi: 10.1191/1478088706qp063oa]
32. Holtzblatt K, Beyer H. Contextual Design: Design for Life. Second Edition. Cambridge, MA: Elsevier; 2017.
33. Broekhuis M, van Velsen L, Hermens H. Assessing usability of eHealth technology: a comparison of usability benchmarking instruments. Int J Med Inform 2019;128:24-31. [doi: 10.1016/j.ijmedinf.2019.05.001] [Medline: 31160008]
34. Broekhuis M, van Velsen L. Improving usability benchmarking for the eHealth domain: the development of the eHealth Usability Benchmarking Instrument (HUBBI). PLoS One 2022;17(2):e0262036 [FREE Full text] [doi: 10.1371/journal.pone.0262036] [Medline: 35176033]
35. Greenaway KH, Kalokerinos EK, Williams LA. Context is everything (in emotion research). Soc Personal Psychol Compass 2018;12(6):e12393. [doi: 10.1111/spc3.12393]
36. Wang T, Giunti G, Melles M, Goossens R. Digital patient experience: umbrella systematic review. J Med Internet Res 2022;24(8):e37952 [FREE Full text] [doi: 10.2196/37952] [Medline: 35925651]

Abbreviations

DPEP: digital patient engagement platform
PX: patient experience

XSL•FO
RenderX

UMUX: Usability Metric for User Experience
UX: user experience

Edited by P Santana-Mancilla; submitted 16.10.23; peer-reviewed by F Wang, H Shen; comments to author 29.12.23; revised version received 21.02.24; accepted 28.03.24; published 14.05.24.

Please cite as:
Aspelund A, Valkonen P, Viitanen J, Rauta V
Designing for Improved Patient Experiences in Home Dialysis: Usability and User Experience Findings From User-Based Evaluation Study With Patients With Chronic Conditions
JMIR Hum Factors 2024;11:e53691
URL: https://humanfactors.jmir.org/2024/1/e53691
doi:10.2196/53691
PMID:38743476

©Anna Aspelund, Paula Valkonen, Johanna Viitanen, Virpi Rauta. Originally published in JMIR Human Factors (https://humanfactors.jmir.org), 14.05.2024. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on https://humanfactors.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Enabling Health Information Recommendation Using Crowdsourced Refinement in Web-Based Health Information Applications: User-Centered Design Approach and EndoZone Informatics Case Study

Wenhao Li[1,2], PhD; Rebecca O'Hara[3], PhD; M Louise Hull[3], PhD, CREI; Helen Slater[4], PhD; Diksha Sirohi[3], MPH, MBA; Melissa A Parker[5], MPhil, RN, RM; Niranjan Bidargaddi[1], PhD

[1]College of Medicine and Public Health, Flinders University, Clovelly Park, Australia
[2]School of Information Science and Engineering, Shandong Normal University, Jinan, China
[3]Robinson Research Institute, Adelaide Medical School, University of Adelaide, Adelaide, Australia
[4]Curtin School of Allied Health, Curtin University, Perth, Australia
[5]Canberra Endometriosis Centre, Centenary Hospital for Women and Children, Canberra, Australia

Corresponding Author:
Niranjan Bidargaddi, PhD
College of Medicine and Public Health
Flinders University
1284 South Road
Clovelly Park, 5042
Australia
Phone: 61 423416543
Email: niranjan.bidargaddi@flinders.edu.au

Abstract

Background: In the digital age, search engines and social media platforms are primary sources for health information, yet their commercial interests–focused algorithms often prioritize irrelevant content. Web-based health applications by reputable sources offer a solution to circumvent these biased algorithms. Despite this advantage, there remains a significant gap in research on the effective integration of content-ranking algorithms within these specialized health applications to ensure the delivery of personalized and relevant health information.

Objective: This study introduces a generic methodology designed to facilitate the development and implementation of health information recommendation features within web-based health applications.

Methods: We detail our proposed methodology, covering conceptual foundation and practical considerations through the stages of design, development, operation, review, and optimization in the software development life cycle. Using a case study, we demonstrate the practical application of the proposed methodology through the implementation of recommendation functionalities in the EndoZone platform, a platform dedicated to providing targeted health information on endometriosis.

Results: Application of the proposed methodology in the EndoZone platform led to the creation of a tailored health information recommendation system known as EndoZone Informatics. Feedback from EndoZone stakeholders as well as insights from the implementation process validate the methodology's utility in enabling advanced recommendation features in health information applications. Preliminary assessments indicate that the system successfully delivers personalized content, adeptly incorporates user feedback, and exhibits considerable flexibility in adjusting its recommendation logic. While certain project-specific design flaws were not caught in the initial stages, these issues were subsequently identified and rectified in the review and optimization stages.

Conclusions: We propose a generic methodology to guide the design and implementation of health information recommendation functionality within web-based health information applications. By harnessing user characteristics and feedback for content ranking, this methodology enables the creation of personalized recommendations that align with individual user needs within trusted health applications. The successful application of our methodology in the development of EndoZone Informatics marks a significant progress toward personalized health information delivery at scale, tailored to the specific needs of users.

XSL•FO
RenderX

(JMIR Hum Factors 2024;11:e52027) doi:10.2196/52027

KEYWORDS

information recommendation; crowdsourcing; health informatics; digital health; endometriosis

Introduction

Background

Members of the general public predominantly resort to search engines such as Google or social media platforms such as Facebook, YouTube, and TikTok as their initial source of health information [1-7]. These platforms use intricate recommendation algorithms to curate the information made accessible to users [8]. The algorithms are designed to rank information based on certain criteria, presenting it in the order of the ranking score. However, the underlying architecture of these ranking systems is by default crafted with commercial intent as opposed to health-centered intent. As a result, information that entices interactions that lead to increased revenue, such as more time spent on the platform or increased traffic to advertisements, gets ranked more prominently. Meanwhile, the information that accurately reflects people's medical needs is buried under large amounts of unrelated articles and posts and becomes difficult to find [9,10].

As the preference for web-based information seeking continues to grow, the development of web-based health information applications by trusted sources has become increasingly popular [11,12]. Through these applications (eg, websites or mobile apps) [13], people can readily access a wealth of health information generated by trusted sources. These interactions present an opportunity to shape an alternative ranking architecture for recommending web-based health content, one that is grounded in health outcomes. The information curated by these trusted platforms is considered superior in quality. Using user behavior after content access to rank health information could pave the way for more effective algorithms. This improved method could be integrated into search engine and social media algorithms through regulatory measures, challenging the current prioritization of web-based health content.

The existing body of research lacks comprehensive guidance on integrating content-ranking algorithms into applications centered around health information delivery. In this paper, we outline a generic methodology to guide the design and implementation of health information recommendation functionality within web-based health information applications. In this methodology, the health information recommendation interface and logic are co-designed with medical experts and application users such as patients and their supporters. This ensures the credibility of the health information provided, as well as accurate reflection of users' preference when interacting with the application. The health information recommended to users is ranked and presented using crowdsourcing technology based on feedback from users who have similar demographic and medical profiles. This ensures that health information can be delivered to people according to their situations and needs. The methodology can be easily integrated into new or existing health information applications. By implementing this ranked health information recommendation feature, we foresee improvements in user experience (UX) and the relevance of health information provided.

This methodology for enabling health information recommendation was first formulated based on our experience and expertise in informatics system development and implementation. It was then further refined and validated through the process of designing and implementing the informatics features of a medical information platform named EndoZone [14]. The platform is funded by the Australian government and Jean Hailes for Women's Health and provides evidence-based information to address symptoms and strategies for managing endometriosis. We illustrate the applicability of the methodology through its application in the EndoZone platform to enable its tailored health information recommendation system known as EndoZone Informatics. The implementation process shows that the methodology is practical for enabling information recommendation functionalities for web-based health information applications that have targeted health content–sharing requirements. Early data show that the solution built using this methodology is effective in reflecting users' feedback and providing highly personalized information recommendations and is also highly flexible in adjusting information recommendation logic. It has also been observed that the design of the user engagement process and user interface (UI) is highly relevant to the rate of users providing feedback and hence can affect the outcome of an information recommendation solution significantly.

The aim of this paper was to outline a generic methodology to guide the design and implementation of health information recommendation functionality within web-based health information applications and demonstrate its application in designing and implementing the informatics features of the EndoZone health information platform.

Related Research

A substantial amount of the articles and videos recommended by search engines and social media platforms have quality issues. They may contain biased content, are not comprehensive enough to cover the topic, are not evidence based, and provide limited coverage or content irrelevant to the topic [5,6,9,10,15]. A review by Osman et al [6] highlighted that >40% of the videos on YouTube on lumbar discectomy, cardiopulmonary resuscitation, and stroke are not useful, while more than half of the videos about vaccination as well as phototherapy and excimer laser treatment for psoriasis reflect bias due to commercial interests. A study assessing the quality of diabetes-related content on TikTok found that the quality of the content varies significantly depending on the types of creators and does not fully meet the health information needs of patients [5]. From billions of web pages and videos on the internet, commercial recommendation algorithms of search engines and social media platforms show those with the highest rank first,

XSL•FO
RenderX

where the ranking criteria often have nothing to do with whether the content could meet people's medical needs [16]. To obtain a higher rank, which leads to a higher visibility rate and eventually a better commercial outcome, billions of dollars have been invested by companies for search engine optimization [8]. This compounds the situation because trusted health information sources such as research organizations and noncommercial health organizations often do not have the financial capacity to compete with commercial companies. As a result, the recommendations made by search engines and social media platforms lead people to unrelated articles, commercial advertisements, or even misinformation. As people generally lack the skills and experience to evaluate the accuracy of the information they are recommended [17], incorrect and harmful medical decisions could be made.

In comparison, web-based health information applications developed by trusted sources such as governments, credited health organizations, and universities provide health information with criteria that people value, such as trustworthiness, expertise, and objectivity [18]. In recent years, many of these applications have been developed globally to bypass the information recommendation algorithms of search engines and social media platforms [11,12,19,20]. Several applications contain mechanisms that provide personalized recommendations of nutritional information, medications, treatment plans, diagnoses or disease predictions, physical activities, or other health care services, based on users' profiles and inputs [21]. However, these recommendation features have not yet been applied extensively in health informatics and medical scenarios [22] and are typically created on an app-to-app basis, targeting a specific disease or recommendation context [12].

The lack of effective information recommendation functionality can be eliminated by enabling health information recommendation capability at scale. Many web-based health information applications could apply similar methodologies in design, development, and evaluation in terms of health information recommendation functionality due to their similarities in context, purpose, and category of recommended items. Tran et al [12] summarized 4 basic recommendation techniques: collaborative filtering, content-based filtering, knowledge-based recommendation, and a hybrid recommendation that combines these 3 techniques. In terms of evaluating the recommendation quality and the effectiveness of the recommendation mechanism, users' feedback is considered to be a major quality criterion [23]. Crowdsourcing technology has been applied in health care and has proven to be an effective approach to collecting retrospective data, such as user feedback, from a large number of dispersed participants [24]. With the development of health informatics technology and current trends of population preferences toward seeking information on the web, the use of crowdsourcing technologies for validating the effectiveness of health information recommendations is promising.

Methods

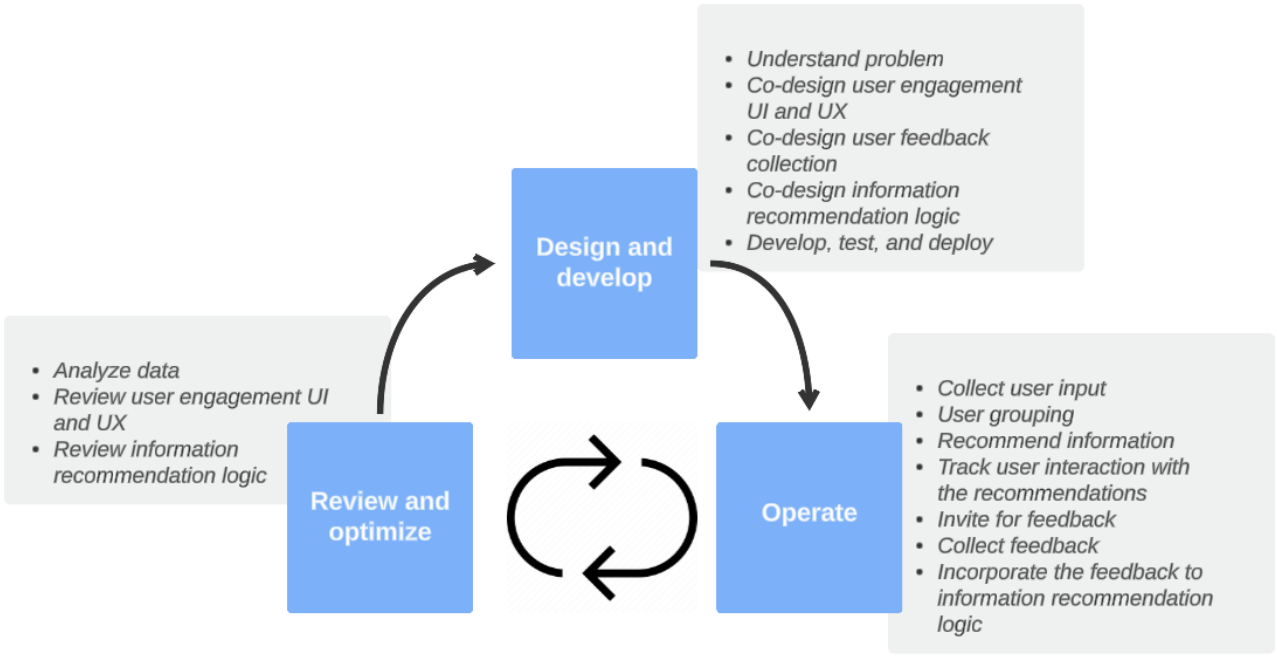
Overview

In this section, we present the concept of the methodology as well as implementation-related design, including software components, software development, and maintenance considerations, during 2 different implementation phases.

The methodology for enabling health information recommendation functionalities involves medical experts, researchers or data analysts, software developers, designers, and users of the web-based health information application. As shown in Figure 1, the methodology consists of 3 stages: *design and develop*, *operate*, and *review andoptimize*. At a high level, the methodology can be summarized thus: first, co-design and codevelop the information recommendation solution; second, recommend information to, and collect feedback from, users to improve the recommendation logic; and third and last, periodically review the statistical data to identify issues and continually adjust the solution.

XSL•FO
RenderX

Figure 1. The information recommendation solution life cycle. UI: user interface; UX: user experience.



Design and Develop

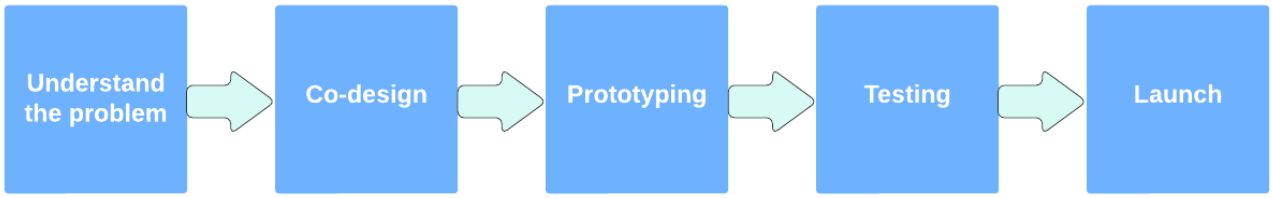
Overview

The implementation of the health information recommendation functionalities in the web-based health information application starts with the design and development of a solution that meets the specific requirements of the application. The design and development process adopts a human-centered design thinking model [25,26]. It considers the need of users to be the main factor that drives the design decision-making process. Figure 2 shows the design and development process for the methodology.

The solution designer first needs to understand the problem and develop detailed requirements for the information recommendation functionalities. To engage in the cocreation process, users can be invited to participate in observation sessions or interview sessions to explore the key challenges and their needs. During these sessions, the questions to be answered could include *What problem is the application trying to solve? What is the status of the application (launched/unbuilt)? What information will be recommended? How is the information expected to be recommended? What data are available for the recommendation logic to be based on?* and *If the application already exists, what does it look like and how is the information recommendation component expected to be integrated?*

Once all requirements have been clearly defined, a series of co-design or ideate sessions are carried out by the solution designer and the medical experts who have comprehensive domain knowledge about the condition or disease the application covers. The co-design process aims to deliver several outcomes, as outlined in the following subsections.

Figure 2. The design and development process.



The Engagement Process

This is also called UI or UX, that is, the UI and the process that the user interoperates with the application's information recommendation functionalities. It will need to be co-designed with medical experts to fully consider users' medical needs.

The Feedback Collection Method

This includes the feedback questions to be asked and the form of questions to be delivered, and it needs to be co-designed with the medical experts as well as the researchers or data analysts, making sure that good UXs and the data collected can properly serve the purpose of the feedback collection.

The Information Recommendation Logic

This specifies how health information that is recommended to users can be realized via different data structures and algorithms. On the basis of the EndoZone Informatics example that we will present later, the recommendation logic could include things such as a list of expert-verified information, a set of rules for information recommendation, an algorithm for user grouping, an algorithm for feedback analysis, an algorithm for feedback incorporation, and an algorithm for information

XSL•FO
RenderX

recommendation. In the co-design process for this deliverable, medical experts should be closely involved in the design of all included components, providing insights that are as detailed as possible and making sure that the recommendations are appropriate (ie, evidence based) and meet users' medical needs. Specifically, the list of expert-verified information and the set of rules for information recommendation should be based on medical experts' input and available research data. Taking a rule in EndoZone as an example, a recommendation of yoga as a self-management strategy is made for a user who has severe pelvic pain and does not experience heavy bleeding during menstruation. In addition to medical experts, researchers or data analysts should also participate in the co-design process, making sure that the algorithms are correctly designed.

The Develop, Test, and Deploy Processes

To ensure that the design fully reflects the users' needs while fully considering the complexity and professionalism of the design activities, a smaller group of user representatives can be invited for consultation, where staged co-design outcomes, as mentioned previously, are sent for review and feedback.

After the co-design process is completed, the solution designer translates the outcomes into system design and architecture specifications, which are then used by software developers to develop, test, and deploy the system. How the develop, test, and deploy processes are carried out depends on the preference of the software development team, where no restrictions are imposed by the methodology. However, it is necessary for stakeholders, including researchers, medical experts, and users, to participate in testing early versions of the solution and provide feedback, where design issues and recommendation logic issues can be identified and resolved in time. The tests can be carried out differently by different stakeholder groups; for example, medical experts and researchers can be asked to test specific features that are closely related to their expertise, while for users of the application, a series of tasks that match their needs and expectations (provided in the initial requirement collection or understanding sessions) can be preset, making sure that their feedback is relevant and targeted.

This *design and develop* stage may be conducted multiple times throughout the lifespan of an information recommendation solution in which the solution is updated to fix issues that are identified and rectified or to incorporate new features.

Operate

Overview

The methodology shifts to the *operate* stage once the information recommendation functionalities are launched. In this stage, the solution performs activities such as recommending health information to users, collecting user feedback on recommendations, and incorporating the feedback into the recommendation logic. The first entry point of the users to the solution should be an event that is related to the content of the application; for example, it can be a click on a button on a web page, an action when using a digital tool, or a click on a link included in an invitation email. Such an event triggers a series of activities to generate a list of recommendations to the user, as outlined in the following subsections.

Collect User Data on Entry

Data stored in the application, such as user account profiles, user input in digital tools, and user browsing history, contain information about the unique circumstance of a user that is needed for personalizing recommendations. When the entry-point event happens, such data are collected for subsequent algorithms.

Group Users

This is an essential step for recommending personalized information to users with different conditions. In this step, the users are grouped by the algorithm for user grouping, based on a set of predefined attributes. Members of a group could have similar demographic and medical profiles, such as condition, age, educational background, symptoms, treatments, and so on.

Recommend Information

Using the data collected on entry as input, an algorithm generates a list of recommended information according to the information recommendation logic. In the algorithm, first, user data collected on entry are checked against the rules for information recommendation; for example, if a user U has symptom S, and the rule R indicates that all users with symptom S will be recommended information I, then I will be recommended to user U. Second, recommendations that match the rules will be ranked according to previous feedback from all users in the same user group. Third and last, the information is shown to the user in the order of the rank, where information with the best feedback (eg, the highest positive feedback rate) is presented first and has a better chance to be viewed.

Track User Interaction

After the information has been recommended, users are likely to read not all but a subset of the recommendations. It is necessary to track which recommendations are read by a user so that in the later step of collecting user feedback to recommendations, questions can be asked effectively. It is assumed that a recommendation has been read by the user if the content has been exposed to the user (eg, the user clicks on a link to an article). Therefore, any interaction that indicates exposure of the information to the user is recorded. Depending on the UI or UX design, recorded interactions could include clicks on a recommendation link or button, the opening of the web page of the recommendation, and so on.

The evaluation of whether the recommendations meet users' medical needs relies on feedback from the users themselves, supported by the power of crowdsourcing. After a certain period of making the recommendations, attempts are made to collect summative feedback from users who may have read the recommendations and potentially carried out practical activities based on the recommendations.

Invite for Feedback

The collection of feedback starts with sending an invitation to the user for participation. If the user has accessed any of the recommendations, an invitation for feedback is sent. Invitations can be sent in the form of oral invitations (eg, telephone invitations or opportunistic face-to-face invitations) or written letter invitations (by post or via email), which will vary from

XSL•FO
RenderX

case to case [27]. The method of sending invitations is determined according to medical experts' suggestions to approach users with specific medical conditions appropriately and maximize the response rate.

Collect Feedback

There are a few ways in which user feedback data can be collected on the web (eg, conducting web-based surveys and allowing user ratings) [28]. Conducting a web-based survey is one of the most popular ways to collect user feedback, is easy to implement, and can meet the requirements of a web-based health information application in many cases. Questions in the survey can be asked from a UX perspective in terms of the helpfulness of the recommendations; for example, questions against a therapy recommendation could include *Did you try this therapy? Did you find the therapy easy to do? How difficult did you find fitting this therapy into your life with your other activities?* and *Did you find this therapy helped in managing your symptoms?* One of the known issues of web-based surveys is the low completion rate [29]. Some strategies to incentivize completion rates can be found in existing studies [30,31].

Incorporate

After a user's feedback is collected and digitized, first, an algorithm for analyzing feedback executes to convert the feedback data into measurable attributes. Second, an algorithm for feedback incorporation deploys these attributes into the information recommendation logic. Depending on the design of the algorithm for feedback incorporation, the outcomes of the incorporation could include an updated set of rules for information recommendation, updated ranks of recommendations, updated descriptions for each recommendation, and so on. After the incorporation process finishes, the user's journey with the information recommendation solution is completed. The updated information recommendation logic will then be applied when other users engage with the information recommendation solution.

Review and Optimize

Overview

As the information recommendation solution operates, system operation data and user engagement data accumulate. Besides using the user interaction data for improving the information recommendation logic in the *operate* stage, an in-depth review and optimization of the solution can be conducted. The purpose is to identify issues based on the analytical outcome of the accumulated operation data set and the experience gained from the continuous operation and maintenance of the solution. Whether the *review andoptimize* stage needs to be carried out depends on several factors, such as the amount of analyzable data accumulated, the urgency of major optimization of the solution to address emerging requirements, and the operation status of the current information recommendation solution. Researchers and software developers need to decide when a formal *review andoptimize* stage is needed. The outcome of the *review andoptimize* stage should include an optimization plan, where detailed redesign and development can be carried out in the following *design and develop* stage.

Analyze Data

User engagement data such as user profile data, data of user interaction with the information recommendation solution, and user feedback as well as system operation data such as operation logs and web-based traffic data are accumulated and of statistical value to the optimization of the information recommendation solution. Depending on the sufficiency of the accumulated data, research questions such as *Do the users engage well with the solution? Is the recommendation solution effective in helping users to find the information that meet their medical needs?* and *Are the recommendations appropriate and suitable for the user to practice?* can be answered, and potential issues in the engagement process and recommendation content can be identified.

Review User Engagement UI and UX

At the *review andoptimize* stage, a retrospective review can be conducted toward the user engagement UI and UX. The review can be based on 2 sources of input: first, it can be based on the researchers' experience gained while continuously operating the user engagement UI and UX. Second, it can be based on user feedback, such as volunteer user group feedback when asked to test and promote the solution. This review could identify design issues in the UI and user engagement process that cause difficulty for users in accessing the features of the solution and the health information they are recommended.

Review Information Recommendation Logic

It is difficult to provide the best configuration to elements of the information recommendation logic and achieve the optimal recommendation outcome during the *design and develop* stage. The reasons include users' composition, uncertainty in user interaction patterns with the application, and a lack of analyzable data. Thus, continuous adjustments to the configuration of algorithms and data structures are needed; for example, grouping attributes and the logic of the algorithm for user grouping, rules for information recommendation, the list of expert-verified information, and the logic for user feedback evaluation can all be fine-tuned to reflect issues identified from the data analysis. In the *review andoptimize* stage, the best configuration for the information recommendation logic should be determined based on testing different configurations. It is most practical to conduct tests on different parts of the information recommendation in parallel with tests in the operate stage of the software to minimize impact to the existing system. To achieve this, a staging infrastructure can be set up, where a mirror copy of the solution can be created for test-related activities.

Implementation Considerations

Overview

This methodology can be adopted for the implementation of health information recommendation functionalities, either with already launched applications or when the application is still under development. In the next 2 subsections, we present implementation-related considerations of the methodology: first, components need to be developed and how these interoperate with other application components is described; and second, a 2-phased implementation strategy that aims to provide the optimal UX is described.

XSL•FO
RenderX

Software Components

The architecture design for the information recommendation solutions could vary vastly due to factors such as user requirements, the software technology stack being applied, the skill sets of developers, and governance restrictions (eg, the General Data Protection Regulation applicable in the European Union). However, when adopting the methodology, logical components for the health information recommendation functionality should be consistent. Figure 3 shows a high-level software component diagram that implements health information recommendation functionalities in a web-based health information application. The diagram consists of 3 sections: components of a typical web-based health information application (Figure 3A), backend components of the information recommendation solution (Figure 3B), and front-end components of the information recommendation solution (Figure 3C).

Figure 3. Software component design. (A) Components of a typical web-based health information application. (B) Backend components of the information recommendation solution. (C) Front-end components of the information recommendation solution.

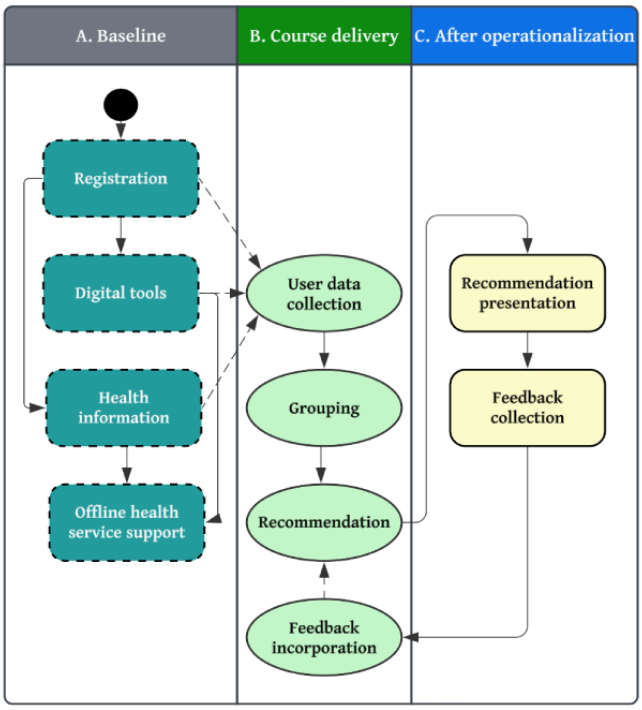


Figure 3A shows that a typical web-based health information application in the form of a website could include web pages of health information; instructions on local offline support; and, optionally, digital tools for certain informatic purposes and a registration component that is often needed by the digital tools. The 4 backend components shown in Figure 3B are needed for an information recommendation, which collects and processes all user input data, makes recommendations, and incorporates user feedback. Depending on the design of the information recommendation logic, the user input data collected could include user account data, health information browsing history, digital tool input, and user feedback on recommendations. The output of the 4 backend components includes updated information recommendation logic and a list of recommendations ranked based on user feedback. The front-end of the information recommendation solution shown in Figure 3C comprises 2 components: one for presenting recommendations and the other for collecting recommendation feedback. Depending on the design of the user engagement process, these 2 components can be either allocated on dedicated web pages or integrated into the web pages of any web-based health information application component.

The software components are designed in a loosely coupled fashion, where all functions and algorithms are independently maintained. Such a design pattern makes the adjustment of the information recommendation logic possible, from fine-tuning to a total replacement of the recommendation model. This feature is critical to a phased implementation of the solution, as will be discussed in the next subsection. One additional advantage of such a design is that it enables the potential of the web-based health information application to become a test bed of information recommendation algorithms, where algorithms can be easily alternated to test performance.

Implementation Phases

When the information recommendation solution is first launched, the number of users is small, and feedback on recommendations is not yet provided. Here, an information recommendation model that relies heavily on crowdsourced data for recommendation evaluation could produce suboptimal recommendations, impacting the UXs with the web-based health information application. To ensure the quality of recommendations before crowdsourced data are sufficiently accumulated, an implementation strategy is applied with the following two phases: (1) an initialization phase, in which crowdsourced data are not yet sufficiently collected, and an initial version of the algorithm for recommending information is used, where the information recommendation logic does not rank the recommendations based on user feedback data; and (2) an execution phase, where crowdsourced data are sufficient, and an execution version of the algorithm for recommending

XSL•FO
RenderX

information is used, where user feedback is incorporated into the information recommendation logic for ranking recommendations based on user feedback data.

The main difference between these 2 versions of the algorithms is their logic in dealing with user feedback. Specifically, in the initial version of the algorithm, the list of recommendations is generated purely based on medical experts' input (ie, a set of predefined rules for information recommendation), whereas in the execution version, the list of recommendations is generated based on medical experts' input and further ranked based on user feedback data. Due to the loosely coupled software design, the algorithm for recommending information can be easily replaced. Researchers and data analysts can decide it is time to replace the algorithm when the amount of user feedback data is sufficient for the execution version of the algorithm to execute effectively.

Ethical Considerations

The development of the platform and analysis of EndoZone data was approved by the University of Adelaide Human Research Ethics Committee (H-2020-013 & H-2023-054). Informed consent was obtained from community members participating in the design and development phase of the EndoZone informatics platform, and all users accessing the tool after it was launched online. The extraction and analysis of de-identified EndoZone platform data for this study was in accordance with the guidelines approved by the ethics committee.

Results

Case Study: EndoZone Informatics

The methodology for enabling health information recommendation functionalities has been successfully applied in the development of the information recommendation functionalities of a co-designed endometriosis information platform called EndoZone [14]. Endometriosis is a chronic condition, where tissue similar to the lining of the uterus develops in places outside the uterus. Symptoms of endometriosis may include pain with menstruation, chronic pelvic pain, fatigue, and subfertility. Globally, it is estimated that endometriosis affects approximately 190 million women and people presumed female at birth [32]. To address the wide-ranging impact of endometriosis, the Australian government and Jean Hailes for Women's Health funded the development of EndoZone to improve knowledge, address symptoms, and provide strategies for managing endometriosis. This platform was designed for people affected by the condition as well as their supporters, such as parents, partners, teachers, and coworkers. The platform was cocreated and developed using the design thinking framework. During the cocreation process of the EndoZone platform, endometriosis community focus groups (n=36) were held to explore the key challenges and needs of the endometriosis community; in addition, a community priorities survey was conducted with 347 community member responses. On the basis of the key priorities identified, it was decided that functionalities would be developed to facilitate interaction and to support people experiencing endometriosis symptoms through the recommendation of strategies based on their symptoms, that is, EndoZone Informatics. The design, development, and implementation of EndoZone Informatics strictly follows the health information recommendation methodology. The solution was co-designed with other components of the EndoZone platform and integrated into the platform in April 2023. The solution is currently fully implemented and operating in the execution phase. In the following subsections, we present the design, development, and implementation of EndoZone Informatics to showcase the practicality of adopting the methodology for the design and implementation of information recommendation functionalities in a web-based health information application.

Design and Develop

The design of EndoZone Informatics was part of the broader platform development process, which follows the broader co-design process of EndoZone. It was designed in consultation with 5 community representatives from endometriosis associations (patients, advocates, and supporters), clinicians (endometriosis or fertility specialist, physiotherapist or pain researcher, and endometriosis nurse), researchers, and 2 health informatics specialists. This involved a series of workshops and meetings to discuss details of the user engagement process as well as a smaller working group with clinicians to develop the initial information recommendation logic. The design was mocked up in consultation with the UI or UX designer and then integrated into the EndoZone platform. The outcome of this co-design process includes the user engagement process and the corresponding UI or UX prototype, a feedback collection method using email invitations and web-based surveys, and the information recommendation logic. Specifically, the information recommendation logic includes a list of 16 expert-verified articles for different endometriosis self-management therapies; a set of 27 rules that match symptoms to the recommendation of therapies (eg, one rule is that if the user experiences severe menstrual cramps, an article on transcutaneous electrical nerve stimulation therapy will be recommended); and algorithms for user grouping, analyzing feedback, feedback incorporation, and recommending information. After the design was ready to be reviewed, a review meeting was carried out for all stakeholders, where feedback on the design outcome was collected for adjustment. After the design outcome was adjusted and agreed upon, the solution designer and the UI or UX designer translated the outcomes to formal UI or UX design and architecture specifications for the development work to be carried out.

The develop, test, and deploy process was carried out using an agile approach, more specifically, the Scrum development process [33], which is preferred by the development team due to the existing software technology stack and developer skill sets. The development progress was regularly reported to, and closely monitored by, the digital health solution transformation experts. One issue that was encountered during this stage was some previously unforeseeable dependencies of the informatics components on several other components of the platform, which caused a 3-month delay in the release date of EndoZone Informatics. However, the development process is in general smooth.

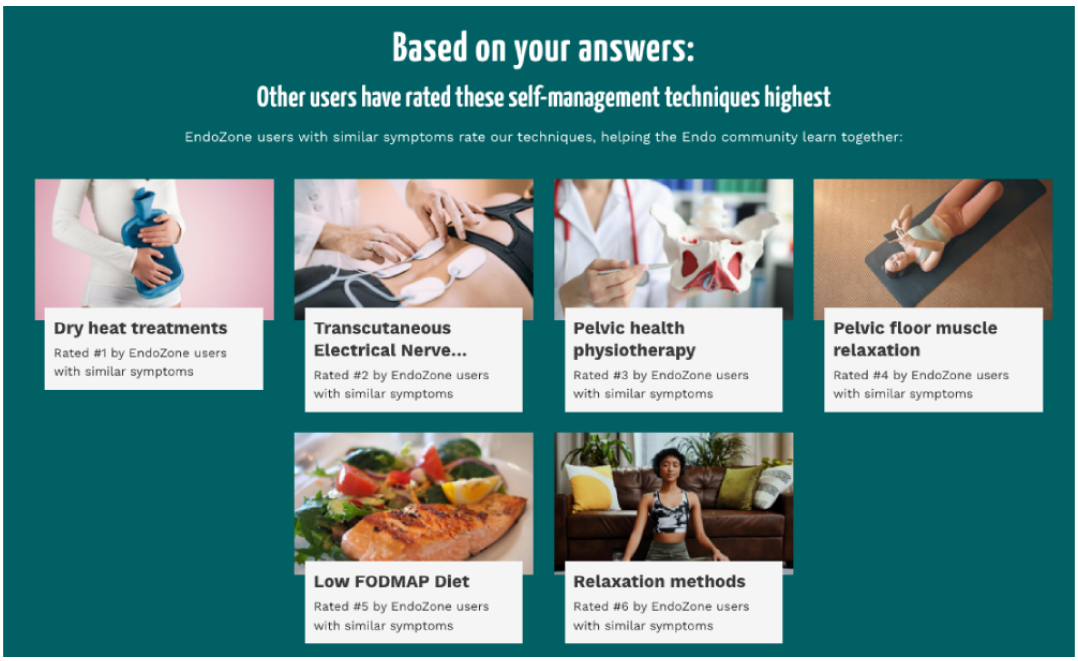
XSL•FO
RenderX

To test EndoZone Informatics, medical and health experts participated in 2 demonstrations of the platform and tested the ready-to-launch version. In all, 9 community users participated in testing the platform's early versions through a beta version with restricted access. Specifically, the user test was conducted after a series of tasks in which user testers were audio and video recorded while they completed the tasks and provided verbal feedback as they were using the platform. They also completed a series of questions related to their feedback on the platform (eg, what they liked, what they did not like, and suggestions for improvement), the usability of the platform, and whether they would recommend the platform to a friend or colleague. The feedback obtained from the user test was then incorporated into the further development of EndoZone Informatics features.

Operate

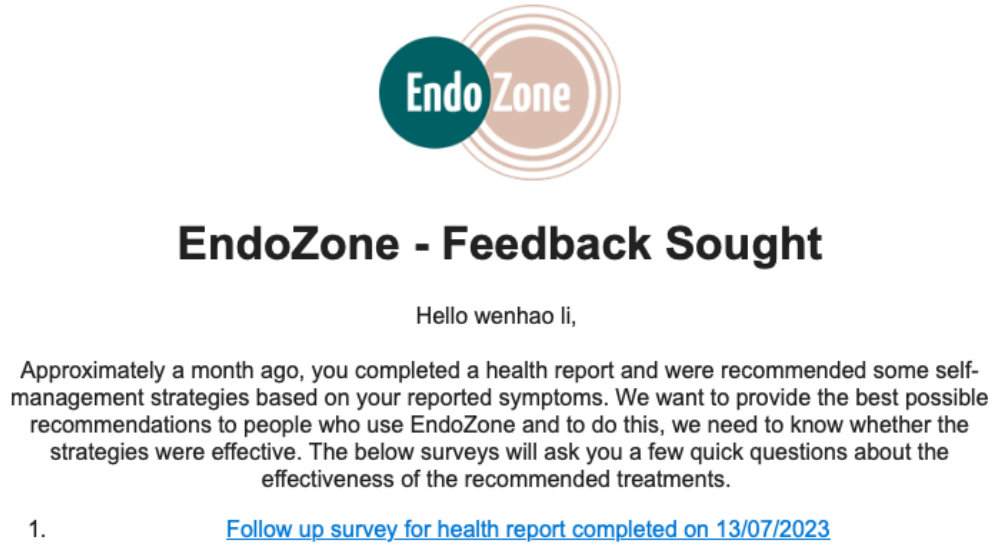
A user starts to engage with the EndoZone information recommendation solution from the submission of a health questionnaire named *My Endo Report*. The questionnaire contains a series of questions related to self-reported endometriosis symptoms and treatments that have been tried to manage symptoms, as well as a brief medical history. After the user has submitted the questionnaire, the backend algorithms are triggered to produce a list of recommended self-management therapies, where the recommendations are presented as part of *My Endo Report* (Figure 4). In EndoZone, the crowdsourced input (ie, user-provided feedback on recommended therapies) is used to determine the order of the recommendations being presented: among users with similar symptoms, therapies that are rated as "more useful" are given a higher rank and shown first in the list of recommendations. The description text of each recommendation contains a ranking to highlight this order. If the user clicks on a recommended therapy, the solution assumes that the user has viewed the content and records the click event. Next, 30 days after the recommendations are made, an email is sent to the user, inviting the user to complete a follow-up survey regarding the recommendations (Figure 5). When the user accepts the invitation, a follow-up survey is generated, containing questions related only to the recommendations that the user has clicked on. For each recommended self-management strategy, the survey contains 10 questions. It asks the user about the usefulness of the strategy, including their feelings after practicing the strategy, the practicality of the strategy, the effectiveness of the strategy in improving their symptoms, and so on. Figure 6 shows an example follow-up survey for the recommendation of pelvic health physiotherapy. Once the user has submitted the follow-up survey, their engagement with the EndoZone information recommendation solution is complete.

Figure 4. Recommended self-management therapies.



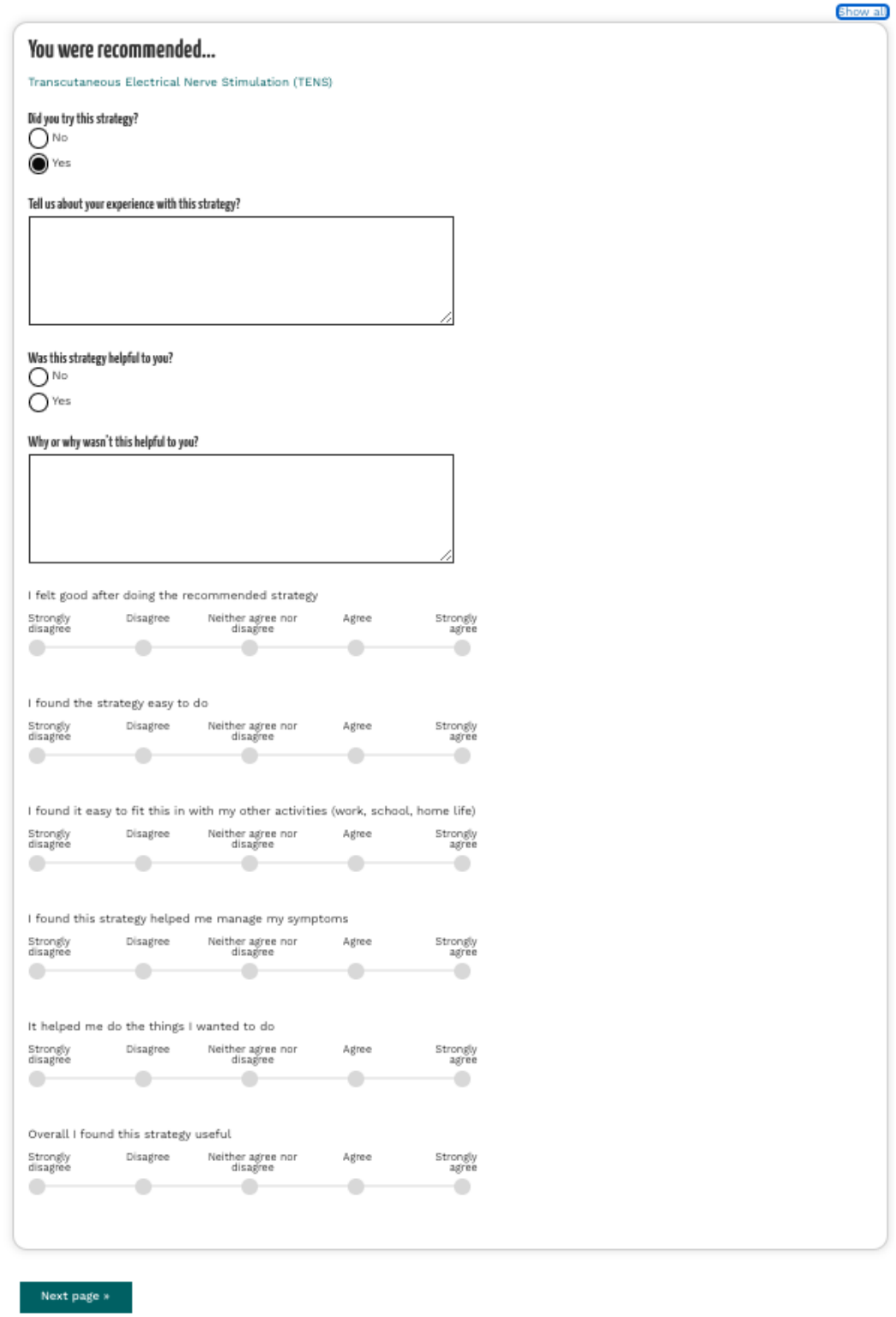
XSL•FO
RenderX

Figure 5. Invitation email to a follow-up survey.



XSL•FO
RenderX

Figure 6. A follow-up survey example.



Software Architecture

The architecture design in the EndoZone information recommendation solution strictly followed the component design shown in Figure 3 but was customized to fit the specific requirements of the application. First, based on the needs of the EndoZone information recommendation logic, the user data collection component only collects user registration data (ie, demographic profile data) and user input to the digital tool (ie, *My Endo Report* submission data). Second, the recommendation presentation component is integrated into the *My Endo Report* summary page of the application as part of the *My Endo Report* outcome.

Deployment wise, based on best practice, the EndoZone information recommendation solution is designed to be cloud based. It operates on cloud-based infrastructure using Amazon Web Services. All recommendation-related components are deployed in the form of microservices using Amazon Web

XSL•FO
RenderX

Services Lambda, where each microservice contains components that are needed for a single application programming interface call. Specifically, the user data collection, grouping, and recommendation components are deployed in 1 microservice. Once the *My Endo Report* questionnaire is submitted, this microservice is called and responds with a list of recommended self-management therapies. The feedback incorporation component is deployed in another microservice. Once the follow-up survey is submitted, this microservice is called to update the information recommendation logic.

Implementation

The implementation process of the EndoZone information recommendation solution followed the 2-phased process. Compared with what is described in the *Overview* subsection in the *Methods* section, an alternative data accumulation approach was conducted in the initialization phase to accelerate the transition to the execution phase. After the platform was launched, a targeted social media campaign on Instagram and Facebook was conducted to promote initial use of the platform. During the campaign, the initial version of the algorithm for recommending information was executed based on the expert-derived set of information recommendation rules that were matched to self-management therapies and symptoms that were indicated in *My Endo Report*. In completing the report, users are contacted via various channels to self-rate how helpful each self-management therapy or strategy was to manage their symptoms using a 3-point scale (“Didn’t work,” “Helped a bit,” and “Helped a lot”).

At the time of reviewing the data, the EndoZone platform had had 57,000 visitors (Google Analytics; February 20, 2024), predominantly from Australia (n=32,000, 56.14%), the United States (n=6000, 10.53%), the United Kingdom (n=5200, 9.12%), and New Zealand (n=5000, 8.77%), of whom 5756 (10.1%) completed *My Endo Report* and submitted it through the platform. User feedback data were aggregated to count the number of reports that indicated that a particular strategy either “Helped a bit” or “Helped a lot.” This feedback was then considered to be the initial rating of therapies on which the execution version of the algorithm could rely; for example, *yoga* was rated by 682 people, of whom 404 (59.24%) rated it as either “Helped a bit” or “Helped a lot.” These feedback data were then manually incorporated, where a rating of “404/682 (59.24%)” was set as the initial rating of the therapy *yoga* for all user groups. A further analysis of the data collected through the platform is being conducted to feed into the next iteration of EndoZone Informatics.

Review and Optimize

The EndoZone information recommendation solution was integrated into the EndoZone platform in April 2023. Tests and feedback from the volunteer group have shown that the overall user engagement process can be carried out well, with a good UX. Meanwhile, based on early data accumulated, several design issues have been identified; for example, the participation rate for providing feedback is lower than expected. We suspect that the UI or UX design could be the major cause for this outcome: first, in the current design, only registered users are invited to complete the follow-up survey (unregistered users cannot be invited because they are not asked for their email address). Currently, most users use the site anonymously, which means that most users of the platform who decided not to create an account in EndoZone are not able to experience the full recommendation functionalities and provide recommendation feedback. Second, the recommendation section is in a relatively inconspicuous position on the *My Endo Report* summary page. This may lead to reduced visibility and hence less user participation. The finding indicates that the design of web pages (UI or UX) is highly relevant to the effectiveness of the solution. It also indicates that the methodology is limited in identifying specific design defects during the initial *design and develop* stage. However, such defects can be addressed in the *review andoptimize* stage, where issues that crop up during the execution of the solution are reviewed. In the context of the EndoZone platform, remedial development work has been planned in the second phase of the project from 2024 to 2026.

Furthermore, the logic for tracking user engagement with recommended therapies (ie, once the article is opened, the recommended therapy is considered to have been read) is not consistent with the industrial standard that large IT companies have applied; for example, in Google Analytics, a user is considered to have engaged with a web page if they stay on the page for >10 seconds [34]. How user engagement is tracked is not defined by the methodology and could vary from case to case. However, in the context of the EndoZone platform, the solution logic does not cause a loss of user feedback data. The impact on the UX (ie, several more survey questions are asked regarding a therapy that the user has not practiced) is limited and can be eliminated by adjusting the questions in the follow-up survey.

Discussion

Outcome

In the previous sections, we have presented a methodology that enables health information recommendation functionalities in web-based health information applications. The concept of the methodology as well as the implementation considerations, including the software component design and the 2-phased implementation process, are described in detail, based on which information recommendation solutions can be created and operationalized. The methodology has been refined and validated through its application to create EndoZone Informatics, that is, the information recommendation solution of an endometriosis information platform named EndoZone. Early data from the execution of the EndoZone Informatics solution shows that using this methodology was effective in recommending medical expert–verified information while incorporating crowdsourced input from users with similar conditions. This methodology helped users to find the information that could be of most use to them. The loosely coupled software component design enabled high flexibility in adjusting the information recommendation model, which makes the 2-phased implementation process easy to carry out.

During the application of the methodology for EndoZone Informatics, we encountered several issues. To recap, first, the dependencies of the information recommendation components

XSL•FO
RenderX

on other components of the web-based health information application caused a 3-month delay in the development progress of EndoZone Informatics. Second, the UI or UX design flaws, such as unregistered users not being able to experience the full recommendation functionalities and *underexposure* of the recommendation section in the *My Endo Report* summary page, have resulted in a lower-than-expected participation rate for providing feedback. These issues reveal a limitation of the methodology, that is, it is not able to address some specific software engineering problems. These issues also show the significance of the *review and optimize* stage, where design and development issues can be identified, and repair plans can be created.

In general, the application of the methodology for designing and implementing EndoZone Informatics is successful. It is a solid step toward enabling personalized information recommendation at scale. The solution indicates a promising approach where personalized health information recommendation can be enabled in all web-based health information applications. Compared with accessing health information via recommendations derived from commercial algorithms of search engines and social media platforms, a health information–access approach provides people with an alternative health information–ranking and –recommendation path, which ranks information based on people's medical needs; provides them with trustworthy, credible, and evidence-based recommendations; and aims for the best health outcomes.

Potential of the Methodology

The methodology is proposed to be applied in web-based health information applications targeting personalized health information recommendations for educational and knowledge-sharing purposes. As showcased by the EndoZone platform, this methodology is applicable and works well for web-based health information applications that share health information such as chronic disease self-management strategies. However, the practicality of applying the methodology in creating solutions for applications that target acute diseases is yet to be proven. Another area for further research is the practicality of applying this methodology for recommending clinical treatments. This requires systematic study of what the impact could be if the methodology was applied for recommending clinical treatments (eg, medication use). What kind of care decisions (safety or risk of harm or relative benefits) need to be considered? What are the ethical issues involved? Answers to these questions are not yet clear.

A promising area for applying the methodology concerns creating solutions for recommending other medical and health services, such as links to local medical experts, health services, advocacy organizations, and related web-based applications [35,36]. Exploring how solutions created by applying this methodology could help in connecting web-based services to local offline services to improve the quality and scope of user support would also be of value.

Another potential application of this methodology is to generate test beds for information recommendation algorithms and their suitability for different medical scenarios. As described in the *Overview* subsection of the *Methods* section, the software component design allows all key logic components to be independently maintained and easily replaced. This feature can be leveraged for new information recommendation algorithms or models to be tested; for example, by applying different information recommendation models and monitoring user interactions under each model, the performance of different information recommendation models can be analyzed.

Conclusions

This study introduces a novel methodology that enriches web-based health applications with personalized information recommendation capabilities. Tested through the development of the EndoZone platform, our approach successfully merges expert knowledge with user insights to provide targeted health information. While we encountered developmental and design challenges, these experiences highlighted the importance of adaptability and continuous refinement. The methodology's potential extends beyond the specific case of EndoZone, offering a scalable solution for tailoring health information across various authoritative health websites, with implications for improving patient education and engagement in a digital era.

Acknowledgments

The authors are grateful for the support and expertise provided by Karl Fradley and Laura Morgan from SRA Information Technology and the team at Brainium Labs, who contributed to the development of EndoZone, which has greatly facilitated the completion of this research. EndoZone was funded by the Australian Government and Jean Hailes for Women's Health.

Conflicts of Interest

None declared.

References

1. Bernard E, Arnould M, Saint-Lary O, Duhot D, Hebbrecht G. Internet use for information seeking in clinical practice: a cross-sectional survey among French general practitioners. Int J Med Inform 2012 Jul;81(7):493-499. [doi: 10.1016/j.ijmedinf.2012.02.001] [Medline: 22425281]
2. Armstrong-Heimsoth A, Johnson ML, McCulley A, Basinger M, Maki K, Davison D. Good googling: a consumer health literacy program empowering parents to find quality health information online. J Consum Health Internet 2017 Jun 12;21(2):111-124. [doi: 10.1080/15398285.2017.1308191]

XSL•FO
RenderX

3. Wong DK, Cheung MK. Online health information seeking and eHealth literacy among patients attending a primary care clinic in Hong Kong: a cross-sectional survey. J Med Internet Res 2019 Mar 27;21(3):e10831. [doi: 10.2196/10831] [Medline: 30916666]
4. Hassan S, Masoud O. Online health information seeking and health literacy among non-medical college students: gender differences. J Public Health (Berl.) 2020 Mar 09;29(6):1267-1273. [doi: 10.1007/s10389-020-01243-w]
5. Kong W, Song S, Zhao YC, Zhu Q, Sha L. TikTok as a health information source: assessment of the quality of information in diabetes-related videos. J Med Internet Res 2021 Sep 01;23(9):e30409. [doi: 10.2196/30409] [Medline: 34468327]
6. Osman W, Mohamed F, Elhassan M, Shoufan A. Is YouTube a reliable source of health-related information? a systematic review. BMC Med Educ 2022 May 19;22(1):382. [doi: 10.1186/s12909-022-03446-z] [Medline: 35590410]
7. Zhao Y, Zhang J. Consumer health information seeking in social media: a literature review. Health Info Libr J 2017 Dec;34(4):268-283. [doi: 10.1111/hir.12192] [Medline: 29045011]
8. Singh Rajawat A, Barhanpurkar K, Goyal SB, Bedi P, Shaw RN, Ghosh A. Efficient deep learning for reforming authentic content searching on big data. In: Proceedings of the Advanced Computing and Intelligent Technologies. 2021 Presented at: ICACIT 2021; March 20-21, 2021; New Delhi, India. [doi: 10.1007/978-981-16-2164-2_26]
9. Hirsch M, Aggarwal S, Barker C, Davis CJ, Duffy JM. Googling endometriosis: a systematic review of information available on the internet. Am J Obstet Gynecol 2017 May;216(5):451-8.e1. [doi: 10.1016/j.ajog.2016.11.1007] [Medline: 27840143]
10. Daraz L, Morrow AS, Ponce OJ, Beuschel B, Farah MH, Katabi A, et al. Can patients trust online health information? a meta-narrative systematic review addressing the quality of health information on the internet. J Gen Intern Med 2019 Sep;34(9):1884-1891. [doi: 10.1007/s11606-019-05109-0] [Medline: 31228051]
11. Martínez-Pérez B, de la Torre-Díez I, López-Coronado M. Mobile health applications for the most prevalent conditions by the World Health Organization: review and analysis. J Med Internet Res 2013 Jun 14;15(6):e120. [doi: 10.2196/jmir.2600] [Medline: 23770578]
12. Tran TN, Felfernig A, Trattner C, Holzinger A. Recommender systems in the healthcare domain: state-of-the-art and research issues. J Intell Inf Syst 2020 Dec 17;57(1):171-201. [doi: 10.1007/s10844-020-00633-6]
13. Baumel A, Kane JM. Examining predictors of real-world user engagement with self-guided ehealth interventions: analysis of mobile apps and websites using a novel dataset. J Med Internet Res 2018 Dec 14;20(12):e11491. [doi: 10.2196/11491] [Medline: 30552077]
14. Welcome to EndoZone. EndoZone. URL: https://www.endozone.com.au [accessed 2023-06-07]
15. Cuan-Baltazar JY, Muñoz-Perez MJ, Robledo-Vega C, Pérez-Zepeda MF, Soto-Vega E. Misinformation of COVID-19 on the internet: infodemiology study. JMIR Public Health Surveill 2020 Apr 09;6(2):e18444. [doi: 10.2196/18444] [Medline: 32250960]
16. Schultheiß S, Lewandowski D. “Outside the industry, nobody knows what we do” SEO as seen by search engine optimizers and content providers. JD 2020 Dec 08;77(2):542-557. [doi: 10.1108/JD-07-2020-0127]
17. Paakkari L, Okan O. COVID-19: health literacy is an underestimated problem. Lancet Public Health 2020 May;5(5):e249-e250. [doi: 10.1016/S2468-2667(20)30086-4] [Medline: 32302535]
18. Sun Y, Zhang Y, Gwizdka J, Trace CB. Consumer evaluation of the quality of online health information: systematic literature review of relevant criteria and indicators. J Med Internet Res 2019 May 02;21(5):e12522. [doi: 10.2196/12522] [Medline: 31045507]
19. Zhao L, Bidargaddi N, Vakulin A, Li W, Luscombe-Marsh N, Benton F, et al. A micro-randomized pilot study to examine the impact of just-in-time nudging on after-dinner snacking in adults with type 2 diabetes: a study protocol. Diabetes Obes Metab 2023 Sep;25(9):2439-2446. [doi: 10.1111/dom.15159] [Medline: 37385960]
20. Zarocostas J. How to fight an infodemic. Lancet 2020 Feb 29;395(10225):676. [doi: 10.1016/S0140-6736(20)30461-X] [Medline: 32113495]
21. Wiesner M, Pfeifer D. Health recommender systems: concepts, requirements, technical basics and challenges. Int J Environ Res Public Health 2014 Mar 03;11(3):2580-2607. [doi: 10.3390/ijerph110302580] [Medline: 24595212]
22. Calero Valdez A, Ziefle M, Verbert K, Felfernig A, Holzinger A. Recommender systems for health informatics: state-of-the-art and future perspectives. In: Machine Learning for Health Informatics. Cham, Switzerland: Springer; 2016.
23. Pu P, Chen L, Hu R. A user-centric evaluation framework for recommender systems. In: Proceedings of the fifth ACM conference on Recommender systems. 2011 Presented at: RecSys '11; October 23-27, 2011; Chicago, IL. [doi: 10.1145/2043932.2043962]
24. McCartney P. Crowdsourcing in healthcare. MCN Am J Matern Child Nurs 2013;38(6):392. [doi: 10.1097/NMC.0b013e3182a41571] [Medline: 24145496]
25. Marko-Holguin M, Cordel SL, van Voorhees BW, Fogel J, Sykes E, Fitzgibbon M, et al. A two-way interactive text messaging application for low-income patients with chronic medical conditions: design-thinking development approach. JMIR Mhealth Uhealth 2019 May 01;7(5):e11833. [doi: 10.2196/11833] [Medline: 31042152]
26. Woods L, Cummings E, Duff J, Walker K. Design thinking for mHealth application co-design to support heart failure self-management. Stud Health Technol Inform 2017;241:97-102. [Medline: 28809190]

XSL•FO
RenderX

27. Bunten A, Porter L, Gold N, Bogle V. A systematic review of factors influencing NHS health check uptake: invitation methods, patient characteristics, and the impact of interventions. BMC Public Health 2020 Jan 21;20(1):93. [doi: 10.1186/s12889-019-7889-4] [Medline: 31964366]
28. Boylan AM, Williams V, Powell J. Online patient feedback: a scoping review and stakeholder consultation to guide health policy. J Health Serv Res Policy 2020 Apr;25(2):122-129. [doi: 10.1177/1355819619870837] [Medline: 31495226]
29. Liu M, Wronski L. Examining completion rates in web surveys via over 25,000 real-world surveys. Soc Sci Comput Rev 2017 Feb 23;36(1):116-124. [doi: 10.1177/0894439317695581]
30. Sammut R, Griscti O, Norman IJ. Strategies to improve response rates to web surveys: a literature review. Int J Nurs Stud 2021 Nov;123:104058. [doi: 10.1016/j.ijnurstu.2021.104058] [Medline: 34454334]
31. Booker QS, Austin JD, Balasubramanian BA. Survey strategies to increase participant response rates in primary care research studies. Fam Pract 2021 Sep 25;38(5):699-702. [doi: 10.1093/fampra/cmab070] [Medline: 34213547]
32. Horne AW, Missmer SA. Pathophysiology, diagnosis, and management of endometriosis. BMJ 2022 Nov 14;379:e070750. [doi: 10.1136/bmj-2022-070750] [Medline: 36375827]
33. Schwaber K. SCRUM development process. In: Sutherland J, Casanave C, Miller J, Patel P, Hollowell G, editors. Business Object Design and Implementation. London, UK: Springer; 1997.
34. What's a good engagement rate in Google Analytics 4? First Page Sage. 2022 Nov 4. URL: https://firstpagesage.com/reports/whats-a-good-engagement-rate-fc/ [accessed 2023-02-03]
35. Hors-Fraile S, Rivera-Romero O, Schneider F, Fernandez-Luque L, Luna-Perejon F, Civit-Balcells A, et al. Analyzing recommender systems for health promotion using a multidisciplinary taxonomy: a scoping review. Int J Med Inform 2018 Jun;114:143-155. [doi: 10.1016/j.ijmedinf.2017.12.018] [Medline: 29331276]
36. Ismail N, Kasim S, Yah Jusoh Y, Hassan R, Alyani A. Medical appointment application. Acta Electron Malays 2017 Nov 01;1(2):05-09. [doi: 10.26480/aem.02.2017.05.09]

Abbreviations

UI: user interface
UX: user experience

Edited by B Price; submitted 21.08.23; peer-reviewed by Y Yang, B Crocker, M Herron; comments to author 12.01.24; revised version received 03.03.24; accepted 31.03.24; published 29.05.24.

Please cite as:
Li W, O'Hara R, Hull ML, Slater H, Sirohi D, Parker MA, Bidargaddi N
Enabling Health Information Recommendation Using Crowdsourced Refinement in Web-Based Health Information Applications: User-Centered Design Approach and EndoZone Informatics Case Study
JMIR Hum Factors 2024;11:e52027
URL: https://humanfactors.jmir.org/2024/1/e52027
doi:10.2196/52027
PMID:38809588

©Wenhao Li, Rebecca O'Hara, M Louise Hull, Helen Slater, Diksha Sirohi, Melissa A Parker, Niranjan Bidargaddi. Originally published in JMIR Human Factors (https://humanfactors.jmir.org), 29.05.2024. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on https://humanfactors.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Participatory Development of an Integrated, eHealth-Supported, Educational Care Pathway (Diabetes Box) for People With Type 2 Diabetes: Development and Usability Study

Daan Leonhard de Frel[1], MD; Mariëlle A Schroijen[2], MD, PhD; Jiska J Aardoom[3,4], PhD; Wesley van Gils[2], RD, MSc; Sasja D Huisman[2], MSc, PhD; Veronica R Janssen[1], MSc, PhD; Anke Versluis[3,4], MSc, PhD; Maaike S Kleinsmann[5], Prof Dr; Douwe E Atsma[1,4,5], MD, Prof Dr; Hanno Pijl[2,3], MD, Prof Dr

[1]Department of Cardiology, Leiden University Medical Center, Leiden, Netherlands

[2]Division of Endocrinology, Department of Internal Medicine, Leiden University Medical Center, Leiden, Netherlands

[3]Department of Public Health and Primary Care, Leiden University Medical Center, Leiden, Netherlands

[4]National eHealth Living Lab, Leiden, Netherlands

[5]Department of Design, Organization and Strategy, Faculty of Industrial Design Engineering, Delft University of Technology, Delft, Netherlands

Corresponding Author:
Daan Leonhard de Frel, MD
Department of Cardiology
Leiden University Medical Center
B4-39g
Albinusdreef 2
Leiden, 2333 ZA
Netherlands
Phone: 31 715262020
Email: d.l.de_frel@lumc.nl

Abstract

Background: Type 2 diabetes (T2D) tremendously affects patient health and health care globally. Changing lifestyle behaviors can help curb the burden of T2D. However, health behavior change is a complex interplay of medical, behavioral, and psychological factors. Personalized lifestyle advice and promotion of self-management can help patients change their health behavior and improve glucose regulation. Digital tools are effective in areas of self-management and have great potential to support patient self-management due to low costs, 24/7 availability, and the option of dynamic automated feedback. To develop successful eHealth solutions, it is important to include stakeholders throughout the development and use a structured approach to guide the development team in planning, coordinating, and executing the development process.

Objective: The aim of this study is to develop an integrated, eHealth-supported, educational care pathway for patients with T2D.

Methods: The educational care pathway was developed using the first 3 phases of the Center for eHealth and Wellbeing Research roadmap: the contextual inquiry, the value specification, and the design phase. Following this roadmap, we used a scoping review about diabetes self-management education and eHealth, past experiences of eHealth practices in our hospital, focus groups with health care professionals (HCPs), and a patient panel to develop a prototype of an educational care pathway. This care pathway is called the Diabetes Box (Leiden University Medical Center) and consists of personalized education, digital educational material, self-measurements of glucose, blood pressure, activity, and sleep, and a smartphone app to bring it all together.

Results: The scoping review highlights the importance of self-management education and the potential of telemonitoring and mobile apps for blood glucose regulation in patients with T2D. Focus groups with HCPs revealed the importance of including all relevant lifestyle factors, using a tailored approach, and using digital consultations. The contextual inquiry led to a set of values that stakeholders found important to include in the educational care pathway. All values were specified in biweekly meetings with key stakeholders, and a prototype was designed. This prototype was evaluated in a patient panel that revealed an overall positive impression of the care pathway but stressed that the number of apps should be restricted to one, that there should be no delay in glucose value visualization, and that insulin use should be incorporated into the app. Both patients and HCPs stressed the importance of direct automated feedback in the Diabetes Box.

XSL•FO
RenderX

Conclusions: After developing the Diabetes Box prototype using the Center for eHealth and Wellbeing Research roadmap, all stakeholders believe that the concept of the Diabetes Box is useful and feasible and that direct automated feedback and education on stress and sleep are essential. A pilot study is planned to assess feasibility, acceptability, and usefulness in more detail.

(JMIR Hum Factors 2024;11:e45055) doi:10.2196/45055

KEYWORDS

diabetes mellitus; type 2; telemedicine; self-management; patient education as topic; activation; glucose regulation; Center for eHealth and Wellbeing Research; CeHRes; type 2 diabetes; tele; patient education; CeHRes roadmap; diabetes; glucose; insulin; education; development; usability; medical; behavioral; psychological; digital consultation; feasibility; endocrinology; endocrine; focus group; dietitian; psychologist; nurse; lifestyle factor; diet; exercise; stress; sleep; cardiovascular disease; heath care professional; mobile phone

Introduction

Background

Around 1 in 11 adults in Europe have diabetes mellitus, and the number of people with diabetes is increasing [1]. Currently, in more than 95% (n>1 million) of diabetes cases, it concerns type 2 diabetes (T2D). In people who are genetically predisposed to diabetes, adverse eating habits, excess body weight, and physical inactivity induce disruption of glucose control [2]. Hence, healthy lifestyle behaviors play a critical role in preventing and managing T2D. Indeed, quitting smoking, being more physically active, eating healthier, and losing weight when overweight can significantly reduce the risk of developing T2D [3,4]. Furthermore, in patients with recently diagnosed T2D, it was demonstrated that dietary lifestyle interventions can lead to persistent diabetes remission after 24 months in 36% (n=149) of patients [5].

Despite the obvious benefits of healthier lifestyles, adherence to healthy lifestyle behaviors in patients with T2D is poor [6-8]. This is alarming, as worse adherence obviously hampers therapeutic efficacy [9]. In Europe, glucose control is inadequate in at least half of the people with T2D. Inadequate glycemic regulation increases the risk of diabetes-related complications and mortality, and it increases medication use and health care costs [10,11]. Immediate action is needed to halt the rising incidence of T2D as well as to decrease the burden of T2D and curb health care costs [1].

A Cochrane review showed that diabetes self-management education (DSME) in people with T2D can improve glucose regulation [12]. In addition, it potentially improves blood pressure and reduces body weight and the requirement for diabetes medication. Yet, another systematic review reported that encouraging patients to play an active role in self-management, so-called patient activation or empowerment, can also improve glucose regulation [13]. Notably, mounting evidence clearly shows that the physiological response to lifestyle change is highly personal [14,15]. Moreover, it seems obvious to suppose that home monitoring of medical and behavioral parameters stimulates and improves self-management. Indeed, integrative monitoring of lifestyle behaviors and physiology using direct action-feedback loops potentially allows for the provision of informative personalized lifestyle advice [16].

Traditionally, DSME is done face-to-face, but digital tools can facilitate health behavior change and significantly improve glucose regulation in patients with T2D [17]. Effective digital tools are self-monitoring (eg, continuous glucose monitoring [CGM]) and telemonitoring by health care professionals (HCPs) [18,19]. Mobile phone apps providing automated feedback can also be effective in improving lifestyle modification and glucose regulation for people with T2D [20]. Indeed, due to low costs as compared to health care consultations and the 24/7 availability of HCPs, mobile phone apps have a lot of potential in diabetes management [21]. Currently available eHealth tools usually focus on one particular lifestyle component or relevant clinical parameter, such as CGM devices or apps that facilitate counting carbohydrates. Examples of digital tools that combine different lifestyle and biometrical parameters to improve self-management and glycemic control exist [22-24]. However, only a few digital tools exist that combine behavioral as well as biological data to provide informed, personalized lifestyle advice to people with T2D [24]. Most of the existing tools are one-size-fits-all lifestyle solutions. Personalized interventions are preferred as clinicians and patients together can choose the treatment plan that contributes most to favorable patient outcomes [25]. Here, we aimed to develop an eHealth-supported educational pathway using integrated behavioral and biological data collected by home monitoring to provide personalized lifestyle advice and promote self-management of people with T2D. Early involvement of stakeholders in the development process of eHealth tools is paramount for successful implementation in health care [26-30]. To assist in the construction of successful eHealth technologies, the Center for eHealth and Wellbeing Research (CeHReS) designed a roadmap to guide eHealth device development, implementation, and evaluation. The CeHReS roadmap consists of 5 phases and emphasizes stakeholder involvement throughout all of these phases [31]. The CeHReS roadmap was used to construct our educational program.

Objectives

Our aim is to empower patients with T2D to manage their disease by developing an integrated, eHealth-supported, blended educational pathway called the Diabetes Box (Leiden University Medical Center). In this paper, we delineate the different phases of the participatory development of the Diabetes Box using the CeHReS roadmap, and the lessons learned are shared.

XSL•FO
RenderX

Methods

Ethical Considerations

The accredited medical research ethics committee Leiden den Haag Delft (MREC registration P21.045) has reviewed the research protocol and gave its approval. The patients participating in the panel provided informed consent for their feedback and input to be used in scientific publication. Input data were deidentified. No compensation was provided for participating in the panel.

CeHReS Roadmap

The CeHReS roadmap was used to guide the development process of the Diabetes Box [31]. The CeHReS roadmap was designed to assist in planning, coordinating, and executing the development process of eHealth tools. The roadmap has a participatory dynamic and consists of 5 intertwined phases and continuous formative evaluation (Figure 1). The first 4 phases (ie, contextual inquiry, value specification, design, and operationalization) of the development process of the Diabetes Box are presented in this paper. The summative evaluation will be performed when the Diabetes Box is launched.

The contextual inquiry is meant to understand the challenges faced by the main stakeholders and how they could be solved. To this end, a literature review was performed. We followed the stages of a scoping review according to the revised Arksey and O'Malley framework [32]. We specified the research question "What diabetes self-management education strategies are being used in regular medical care?" The search strategies combined the terms "diabetes self management education," "technology/telemonitoring/glucose monitoring," and "healthcare/medical care." We searched PubMed and used Google for a broader search. One researcher (DLF) selected the studies and discussed these with a team consisting of 2 endocrinologists, a psychologist, a dietitian, and 2 diabetes nurses, all experienced in the field of DSME. Relevant studies were selected and summarized after which the team discussed the report. To elaborate further on the review of literature, previous experiences with eHealth in our center were evaluated, and important stakeholders were identified and interviewed in focus groups. Previous experiences mainly included technological and practical considerations from implementations of eHealth for patients with myocardial infarction, cardiac surgery, and COVID-19 [33-35]. The main stakeholders were patients, medical specialists, dietitians, psychologists, and diabetes nurses. The latter 4 would later form the development team and partake in the first 2 focus groups.

During the second phase, the value specification, the values gathered in the first phase were translated into (technological) requirements. What problems should the tool solve and how should it work? Weekly meetings with the relevant stakeholders (identified during the contextual inquiry) were used to refine the values and specify the technological requirements of the Diabetes Box.

Using these requirements, prototypes of the Diabetes Box were created during a highly dynamic, iterative, and collaborative design phase. Through biweekly meetings, the development team and stakeholders collaborated closely to ideate, create, and discuss ideas. A panel of patients with T2D gave feedback on the prototype. The entire development team was present on the web during the patient panel. Two members of the team wrote a summary of the recording, after which the recording was deleted. The entire team came together to discuss the outcomes of the patient panel, extract the most important aspects, and set out to change the prototype accordingly. Throughout the development process, the development team looked back on values and knowledge from previous phases to check the integrity of the design. Furthermore, at any point, incoming information could lead to adaptions in the process. This formative evaluation was enabled by constantly involving stakeholders in evaluations and decision-making.

When the design satisfied all stakeholders, the operationalization phase began. During this phase, the Diabetes Box was put into practice. First, a plan was made to implement the newly developed technology into the context defined by the contextual inquiry. The plan was made in close cooperation with the stakeholders to ensure a good fit. Second, the technology is launched.

In the fifth and last phase, the summative evaluation, the tool will be tested in the real world. Currently, the development team is setting up a pilot study to evaluate the feasibility, acceptability, and usability in clinical practice and get an impression of the clinical effects. It is important to note that the technology is quite versatile and adaptable to suit the practical demands of stakeholders as revealed during phase 5. A summary overview of all phases in this study is presented in Figure 2.

XSL•FO
RenderX

Figure 1. Overview of the CeHRes roadmap showing the different phases and formative evaluation (adapted from van Gemert-Pijnen et al [31]). CeHReS: Center for eHealth and Wellbeing Research.

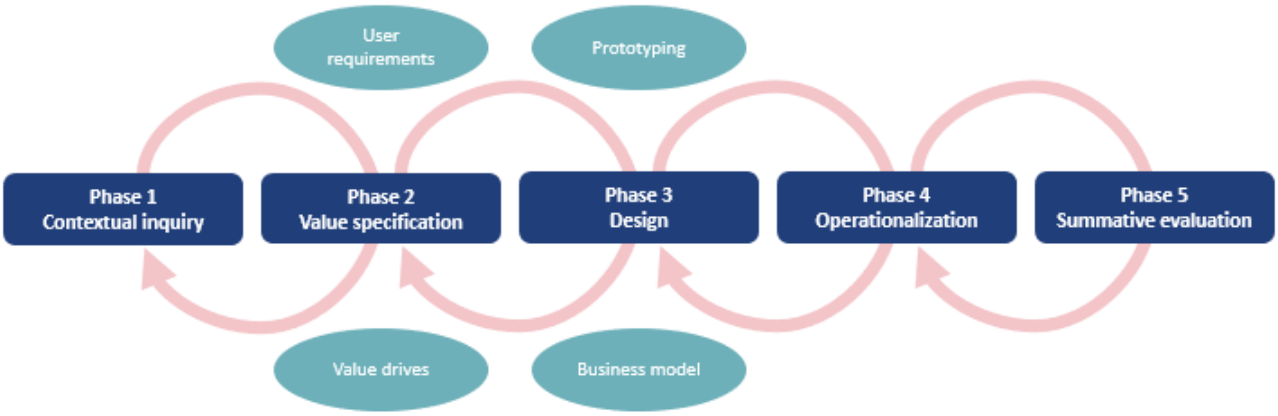
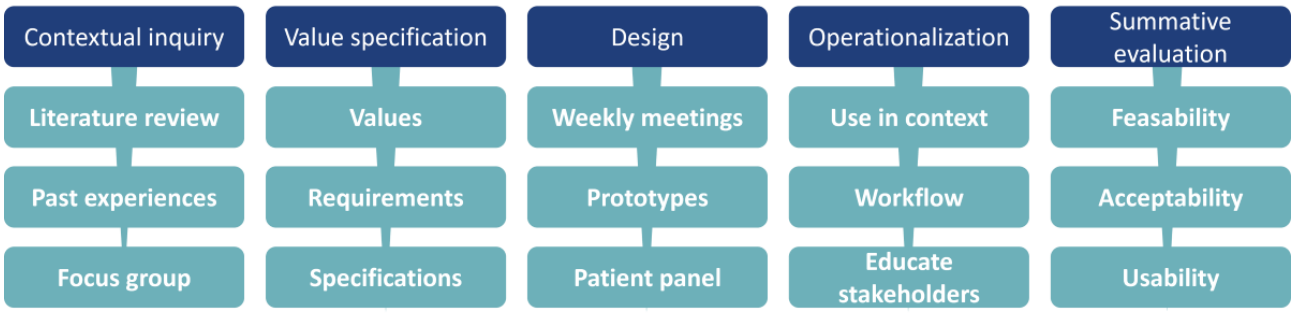


Figure 2. Overview of aspects in every phase. Note that this paper focuses on the contextual inquiry, the value specification, and the design phase [31].



Results

Phase 1: Contextual Inquiry

The contextual inquiry was meant to understand the challenges faced by the main stakeholders and how these challenges could be solved. We used a review of the literature and evaluated past experiences and focus groups with important stakeholders.

Literature Review

The information gathered from the literature review is summarized in Table 1. Studies have outlined several ways to support self-management, which can be categorized as education, monitoring, and modalities. Evidence shows that group-based education about disease pathophysiology, the influence of lifestyle (diet, exercise, stress, and sleep), self-management, and patient activation can improve glucose regulation, reduce body weight, and reduce the need for diabetes medication [4,12,17]. However, HCPs generally feel that they are insufficiently equipped to provide patients with T2D with the insights required to facilitate their health-related behavior change [36-38]. Furthermore, studies suggest that dietician-led lifestyle intervention as compared to interventions led by other HCPs achieves greater weight reductions [39].

Studies on monitoring indicate that self-monitoring (patients monitoring their own health parameters) and telemonitoring (using information technology to monitor patients at a distance) can significantly increase glucose regulation and reduce T2D-related complications [19]. For example, CGM significantly improves glucose regulation and reinforces patient satisfaction [18,48]. In addition, even though activity tracking has ambiguous effects on glucose regulation, it appears to reduce mortality and CVD risk in patients with T2D as well as the incidence of T2D in a general population [51,52]. Furthermore, blood pressure monitoring can decrease systolic blood pressure in patients with T2D when supported by an HCP [53]. To our knowledge, weight monitoring has not been assessed as a stand-alone intervention, but focus on weight can lead to stigma in patients with T2D, potentially leading to increased emotional distress [55].

XSL•FO
RenderX

Table 1. Outcomes of literature review.

Effects	Comments
DSME[a]	
• Self-management ↑[b] • Glucose regulation ↑ [4,12,13,39] • Knowledge or insight ↑ [12] • Body weight ↓[c] [12,39] • Need for medication ↓ [12] • Blood pressure ↓ [12]	• The inclusion of disease pathophysiology contributes to the effect of DSME [12] • Discussing the influence of lifestyle factors (diet, exercise, sleep, and stress) in DSME is decisive for its improvements [4,12,40-46] • Empowerment and patient activation beyond mere education are important in DSME [13] • Digital components of education can be effective [17,39] • The involvement of a dietitian increases the effect on body weight [39]
Telemonitoring	
• Glucose regulation ↑ [19] • Diabetes-related complications ↓ [19]	• Manual input may lead to erroneous input and can lower compliance [47]
CGM[d]	
• Glucose regulation ↑ [18,48] • Patient satisfaction ↑ [49]	• Failing to integrate well-structured education in glucose monitoring can diminish the effects on glucose regulation [50]
Activity tracker	
• Glucose regulation ↑/↓[e] [51] • Incidence T2D[f] ↓ [52] • Mortality ↓ [51] • CVD[g] risk ↓ [51]	• N/A[h]
Blood pressure monitor	
• Systolic blood pressure ↓ [53]	• HCP[i] support increases the effect [53,54]
Weight monitoring	
• Not assessed as a stand-alone	• Focus on weight monitoring and loss can be stigmatizing and lead to increased diabetes-related distress [55]
Mobile apps	
• Glucose regulation ↑ [56] • Monitoring or education ↑ • Lifestyle modification ↑ [20]	• HCP support increases the effect [56-58] • User-friendliness is an important aspect for success [58] • Apps can provide insight into self-management [57]
Dietary journal	
• Inform patients ↑ • Evaluate interventions ↑	• An easier, less time-consuming method would be beneficial to adherence • Photos have equal results as food weighing [59]

[a]DSME: diabetes self-management education.
[b]↑: improves.
[c]↓: deteriorates.
[d]CGM: continuous glucose monitoring.
[e]↑/↓: ambiguous results.
[f]T2D: type 2 diabetes.
[g]CVD: cardiovascular disease.
[h]N/A: not applicable.
[i]HCP: health care professional.

As far as modalities are concerned, mobile phone apps have a lot of potential in T2D management due to low costs, 24/7 availability, and dynamic automated feedback [21]. Evidence points out that mobile phone apps providing lifestyle advice can improve glucose regulation and facilitate lifestyle modification, particularly when they are supported by high-frequency HCP feedback [20,56,57]. Furthermore, keeping electronic dietary records effectively informs patients about the impact of food on glucose levels, but easy-to-use technology is needed [60,61]. For example, taking pictures of meals may be an adequate alternative of time-consuming, labor-intensive recording of dietary components [59].

XSL•FO
RenderX

What Experiences do we Have With Digital Tools Supporting Self-Management?

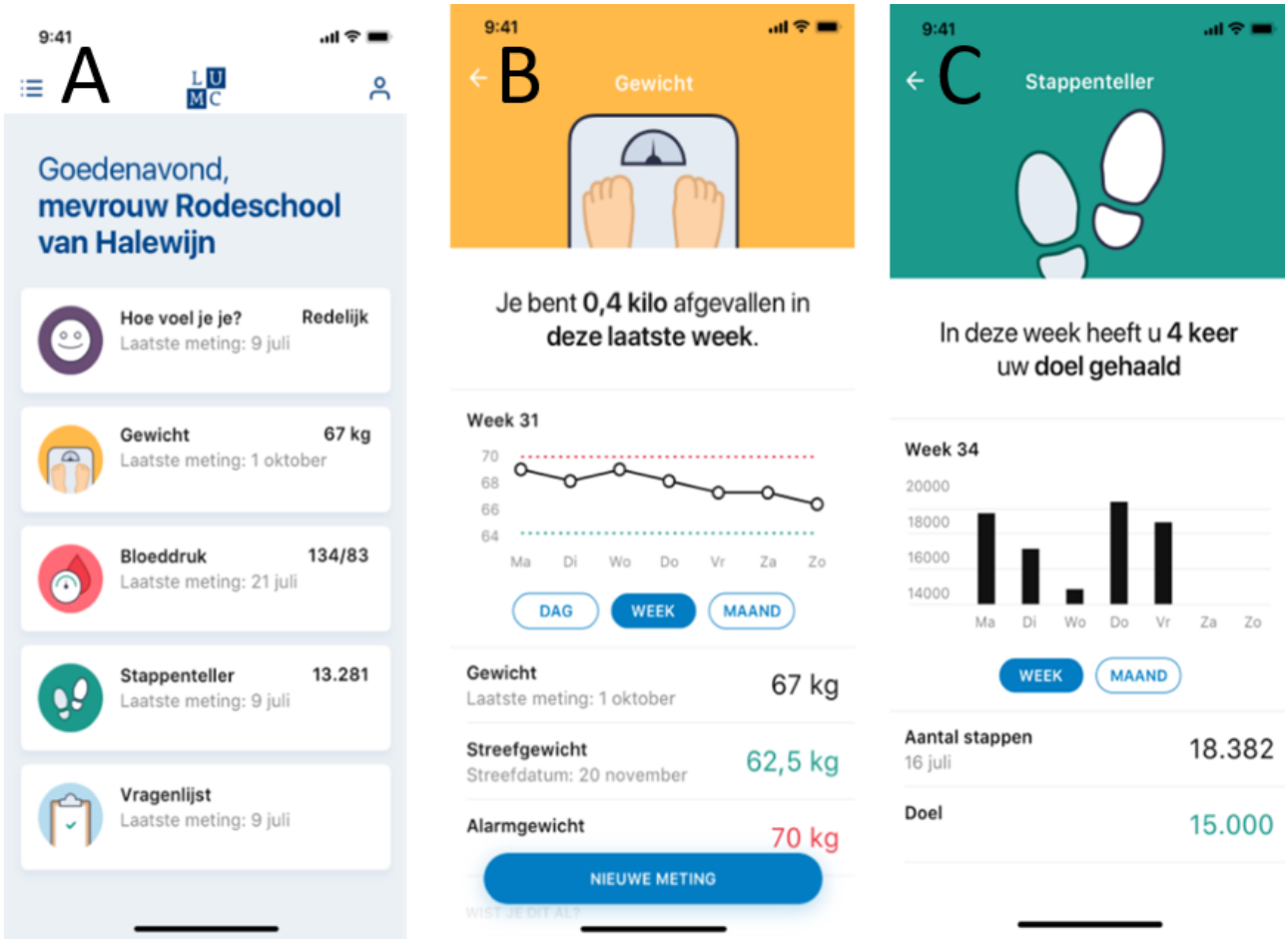
The Box

The Box (Leiden University Medical Center) comprises a set of eHealth tools that aim to improve self-management skills for a specific chronic condition. It includes devices for home monitoring of biological and behavioral parameters relevant to health (eg, glucose concentrations, physical activity, or blood pressure). The data are presented to the patient in a smartphone app called the LUMCCare app (discussed in LUMCCare App section). The data are also sent to the patients' electronic medical records in the hospital to allow evaluation by HCPs. The efficacy and safety of the Box have been examined in the follow-up care of patients with myocardial infarction. We recently reported that patient satisfaction with the Box was equal to regular medical care and that 96% (n=100) of participants appreciated that they could view their health data [62]. Furthermore, the Box has been shown to reduce hospital admissions by effectively surveying clinical symptoms and vital signs at home in patients with COVID-19 [35]. In conclusion, the Box appears to be an effective and appreciated prototype instrument for home monitoring that can be tailored to the health care needs of different conditions.

LUMCCare App

The LUMCCare app (Leiden University Medical Center) is a smartphone app available for Android and iOS. All data collected by the devices in the Box are automatically sent to the LUMCCare app via Bluetooth. The LUMCCare app was codeveloped with people with low health literacy, ensuring a good understanding and usability also in those individuals. Currently, the app was developed in the Dutch language and can display measurements of weight, blood pressure, heart rate, electrocardiograms, steps, temperature, and oxygen saturation (Figure 3). Moreover, users can indicate their level of well-being and provide a brief explanation. HCPs can also send questionnaires to patients through the app. The vast majority of patients with myocardial infarction report intensive and consistent use and high satisfaction with the app [62].

Figure 3. Screenshots of the LUMCCare app from left to right: (A) the home screen showing general well-being, weight, blood pressure, activity, and questionnaires, (B) the weight screen, and (C) the activity screen.



Stakeholders, Current Situation, and Experiences

The main stakeholders included patients, medical specialists, dietitians, psychologists, and diabetes nurses. To evaluate and confirm the findings of our scoping review and past experiences, we organized focus groups with a professor of diabetology, a clinical endocrinologist, a dietician, a psychologist, 2 diabetes nurses, an IT specialist, and a researcher.

According to international guidelines, people with T2D at least annually visit a physician (endocrinologist or general practitioner) and a nurse specialized in diabetes care. These HCPs should educate patients on (the role of lifestyle in) the

XSL•FO
RenderX

pathophysiology of diabetes and on the types and dosing of available medication. Based on patient needs and health parameters, they decide if the patient requires a consult with the dietician or psychologist. All international guidelines advocate lifestyle intervention as a first step in the treatment of T2D. However, health care systems generally lack the means to adequately support patients trying to change deeply engrained habits. This is made exceedingly difficulty by an environment that relentlessly entices them to make unhealthy choices. All HCPs confirmed that continuous home monitoring of subcutaneous glucose concentrations has been a significant advance in supporting and motivating patients with diabetes to enhance their own grip on disease management. The notion that home monitoring of various relevant behavioral and biological parameters and integrating the data to yield personalized feedback would enhance patient empowerment and potentially improve self-management was broadly shared. To these ends, the contents of the Box and LUMCCare app were envisioned to require specific features as further defined in the next stage of development. A summary of the current situation is provided in Textbox 1.

Textbox 1. Current situation of health care for patients with type 2 diabetes (T2D) according to health care professionals (HCPs).

What is going well?

- Knowledge on diet, exercise, stress, and sleep is intermittently conveyed to patients by HCPs.
- All patients see an endocrinologist and diabetes nurse. If deemed necessary, a dietician and a psychologist are available.
- Adequate optimization of medication use.
- Close interdisciplinary collaboration between doctors, dieticians, psychologists, and diabetes nurses in the care for patients with diabetes.
- Health care can be delivered through digital means.

What can be improved?

- Patients' knowledge regarding the influence of lifestyle behaviors (diet, exercise, sleep, and stress) on glucose regulation.
- Activation of patients with T2D to improve lifestyle behaviors.
- Personalized lifestyle advice.
- Focus on personal goal setting.
- Shared decision-making regarding the timing and intensity of consultations with HCPs.
- Home monitoring of relevant parameters.
- Digital group consultations.

Phase 2: Value Specification

After a thorough exploration of the context and potential improvements, the next step was to translate the requirements of eHealth tools that were identified by the HCPs into specific technological properties. First, HCPs emphasized the need to more extensively convey the importance of lifestyle behaviors, including diet, exercise, stress, and sleep, in the control of glucose metabolism and the treatment of T2D. HCPs also stressed that the tool should tailor information and advice to the needs and wishes of the patient and that it should be easy to use for both patients and HCPs. Moreover, it should have features that activate patients to appropriately adapt their lifestyle. Activation was listed as a separate capacity of the eHealth tool. Finally, and importantly, the capability to monitor relevant parameters at home and easy accessibility to collected data for patients and HCPs were defined as prerequisites of an effective tool. This leads to a complete list of values, tool requirements, and tool specifications (Table 2).

XSL•FO
RenderX

Table 2. User perspective, user values, tool requirements, and tool specifications of the Diabetes Box.

Values	Tool requirements	Tool specifications
Provide insight, holistic view	Provide tailored education on the relationship between specific lifestyle factors and glucose regulation	• Include a graph of glucose combined with relevant lifestyle factors (diet, activity, sleep, and stress). • Include education on these topics.
Activate and stimulate	Help stimulate patients to adopt healthy behaviors	• Include goal setting in all aspects of self-management education. • Provide direct behavior–related feedback and education.
Personalized	Tailored to the patient	• Provide a place where patients can monitor their own personal and combined parameters.
24/7 Availability	Rely mostly on apps and e-learnings that are available 24/7	• Provide a digital resource that patients can use in their own time to measure glucose, diet, activity, stress, and sleep. • Provide links for access. • Resources or videos.
Integrated in health care	Integrated in health care	• Add Diabetes Box dashboard to the electronic medical record. • Plan education by HCPs[a] in work hours.
User-friendly	Easy to use, logical, and understandable	• Use B1-level language throughout the tool. • Simplify user interface. • Incorporate dashboard into electronic medical record. • Align education contents with the expertise of HCPs.
Monitoring patients	Monitoring of patient parameters and making them available for patient and HCP	• Provide a place where patients can gain insight into personal parameters of lifestyle factors and glucose. • Provide a dashboard where HCPs can monitor combined patient parameters to provide tailored support.
Low costs	No extra costs for patients, lower costs for health care	• Build on existing app content. • Use group meetings. • Free for patients (insurance covered).
Service desk	Enable a patient support desk	• Two separate phone numbers were provided for difficulties. First, the outpatient clinic number for diabetes-related questions, and second, the Box support desk for technology-related questions.

[a]HCP: health care professional.

Phase 3: Design

Overview

The Diabetes Box was developed using the participatory development method guided by the CeHReS roadmap. The design revolves around making prototypes of the Diabetes Box based on the tool specifications identified in the previous phases and gathering feedback from stakeholders. Our initial prototype was presented to HCPs and shared in a patient panel described below. After feedback from the HCPs, the Diabetes Box comprised digital self-measurement tools, an app, and DSME in the form of consultations and instructive videos (Figure 4). The tools included a continuous glucose monitor (Abbott Freestyle Libre), a sleep or activity tracker (Withings HR Steel), and a blood pressure monitor (Withings BPM Connect). The data collected were presented in the LUMCCare app. Subjective stress could also be registered, and food intake could be monitored by pictures taken of all that was consumed (Figure 5). All data were easily visualized in daily, weekly, and monthly overviews. The data of diet, activity, sleep, and subjective stress could also be plotted on the continuous glucose graph to provide insight into the relation between lifestyle factors and glucose regulation. An expert-led educational program was developed to further promote knowledge of the relationships between various components of lifestyle and glucose control. The overarching goal of the Diabetes Box is to empower patients and facilitate self-management of their disease. The program combines knowledge from routine diabetes care provided by dietitians, psychologists, endocrinologists, and specialized nurses. All educational material was developed, aiming to promote patient self-management. Therefore, multiple behavior change techniques were included in the development of the Diabetes Box. These included information provision, goal setting, action planning, self-monitoring, feedback provision, social comparison, and motivational interviewing. In Multimedia Appendix 1, we provide a list of behavior change techniques as described by Michie et al [63], including a description of the app in the Diabetes Box. There were nine 3- to 5-minute educational videos combining live feed and animations. The topics of these videos were an introduction to the Diabetes Box, the pathophysiology of diabetes, CGM, diet, exercise, sleep, stress, goal setting, and self-management. The educational

XSL•FO
RenderX

program also entailed five consultations: (1) a group consultation introducing the digital tools, (2) an individual consultation focusing on diet, (3) a group consultation regarding diet and exercise, (4) a group consultation regarding sleep and stress, and (5) an individual consultation to evaluate, conclude, and set up future goals. Goal setting and patient activation were present in all videos and all educational consultations. The consultations lasted 45-90 minutes. Prior to each consultation, participants were asked to watch 1 or 2 videos.

Figure 4. Overview of the educational pathway for patients with T2D using the Diabetes Box.

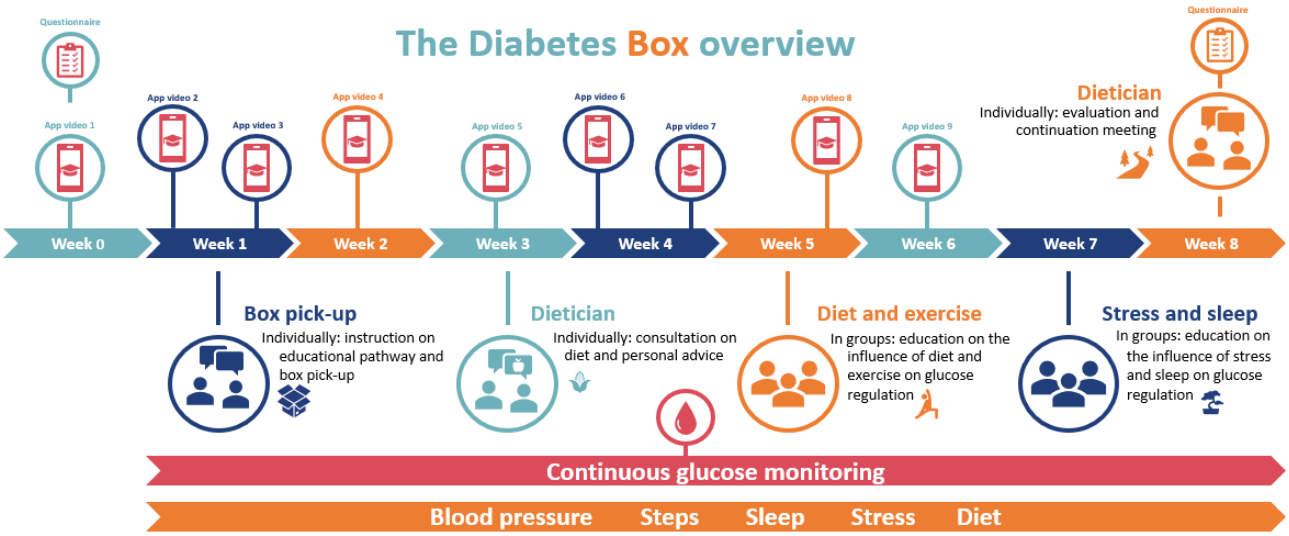
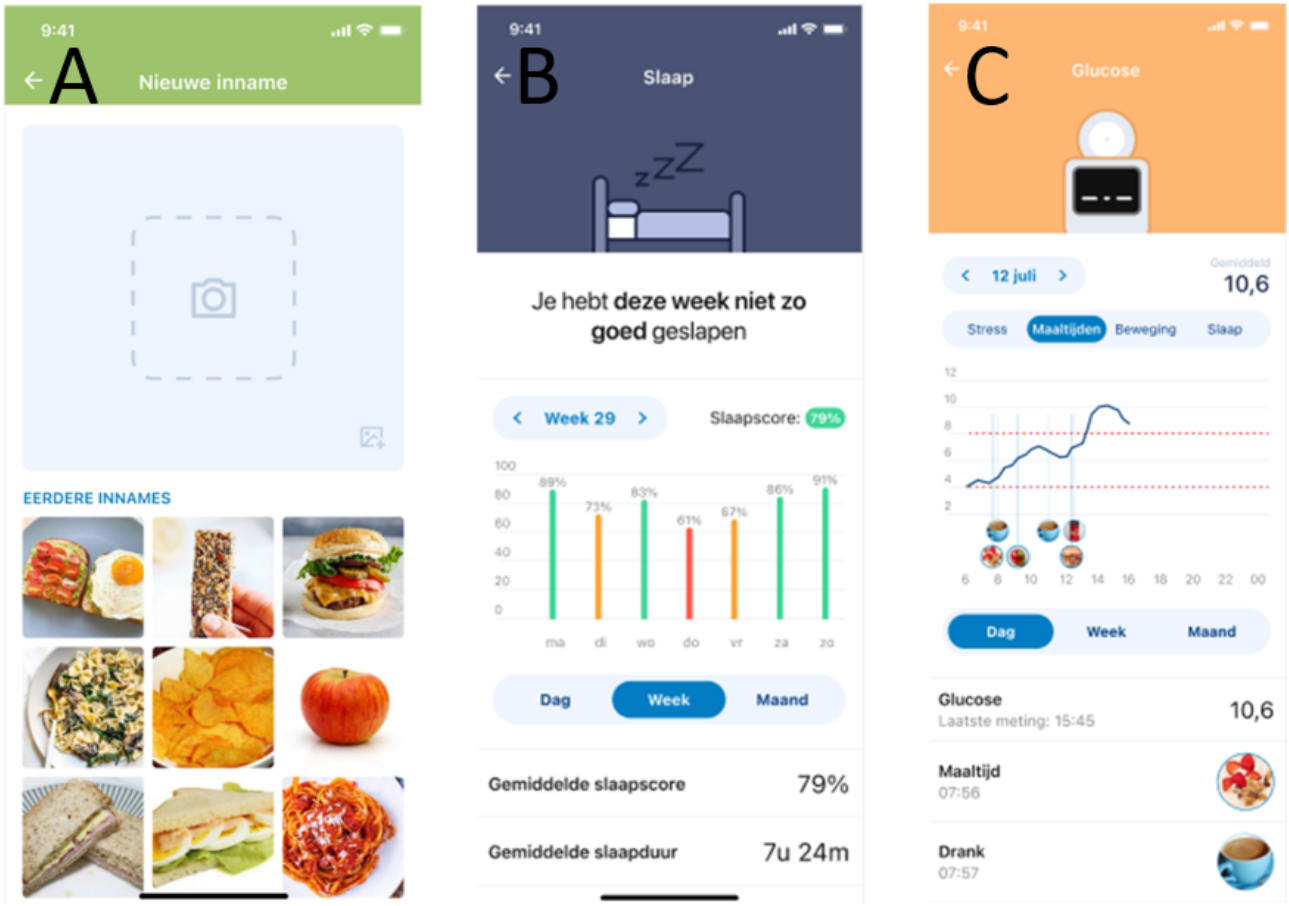


Figure 5. Screenshots of the new modalities of the LUMCCare app from left to right: (A) the screen to register a new intake, either food or drink, with a photo; (B) the sleep screen showing an average score based on estimated duration, interruptions, and regularity; and (C) the glucose screen combined with the diet screen showing the photos of intakes in the glucose graph.



XSL•FO
RenderX

Patient Panel

The prototype of the Box was shared with people with T2D to gather feedback on the preliminary design. Due to COVID-19 measures, only 4 people with diabetes were present during the session. The session took 90 minutes. The concept of the Box program was explained, 2 of the 9 educational videos were shown, the app and its functionalities were demonstrated, and people could try out the eHealth tools. The panel was asked questions covering 4 domains: general opinion, contents and clarity of the educational videos, functionality of eHealth instruments, and usability of the app. Overall, the evaluation was positive, while several potential improvements were suggested (Table 3).

Based on the feedback provided by the patient panel, multiple adaptations were made. First, a leaflet was added to the Box, explaining the flow of the program and anticipated time investment from the side of patients in more detail. Second, a web page was made, displaying the videos accompanied by instructions on when to watch which video. Third, a handout was made with detailing information about the different types of diabetes medications, their uses, and their common side effects. Fourth, as patients preferred a mix of consultation types, 2 consultations were planned online and 3 face-to-face. Fifth, the video of sleep was cut into 2 halves of 3 minutes to prevent viewers from quitting halfway through. Last, the LUMCCare was further developed to also accommodate insulin registration and other activities than steps (eg, cycling or swimming).

Table 3. Outcomes of patient panel.

Comments	Actions
General opinion	
• Positive first impression • Useful and feasible • Questions about total time investment, duration of the education, and loan of the devices • Advise to restrict the number of needed apps to one	• Informative material was made for participants of the Diabetes Box addressing the expected time investment, duration of the pathway, and the fact that participants can keep the devices. • Technologically we still need 2 apps. However, 1 only needs to run on the background and does not have to be opened.
Education	
• Positive • Duration and frequency seem feasible • Half preferred online (travel distance and comfort of home) and half preferred live (connection with HCP[a]) • Early evening is the best time • Videos are appreciated (up to 3 minutes) • A clear overview of videos and when to watch them is needed • Advice in education has to be consistent • Extra attention for diabetes medication, not all glucose levels can be related to behavior, correct use of the CGM[b]	• Hybrid pathway, part of consultations live, part online. • The video of sleep was cut into 2 parts. • A web page was made with an overview of all videos explaining when to watch which. • An extra handout was made about diabetes medication. • Attention to explaining glucose levels was added to the information.
Devices or measurements	
• Doable, clear, and easy to use • Frequency of measurement was regarded positively • Activities other than steps would be great • A (3 hours) delay in showing glucose values was deemed very impractical	• Added activity tracking other than steps in the app. • Focused on displaying real-time glucose data in the app.
LUMCCare app	
• Positive about layout and readability • Stress measurement and diet photos were deemed useful • Diet photos were deemed confronting in a helpful way • Incorporate insulin use in the app	• A functionality to register insulin use was added to the LUMCCare app.

[a]HCP: health care professional.

[b]CGM: continuous glucose monitoring.

Phases 4 and 5: Operationalization and Summative Evaluation

Operationalization involves the introduction of eHealth technology into practice. To test our design, we are currently planning a pilot study in 32 people with T2D to assess the feasibility, acceptability, and usability of the Box in clinical practice. Secondary objectives are evaluation of time in range and perceived learning. The study duration will be 2 months (as the concept Box program lasts 2 months). Participants will fill out a questionnaire before and after the study, and they will be interviewed about their experience as well. Patient satisfaction, user-friendliness of Box components, added value of the program in terms of disease management, and eventual use of the help desk will be evaluated. Consultation attendance, the use of eHealth tools and apps, and eventual replacement of glucose monitors will be registered. HCPs will be asked for

XSL•FO
RenderX

their opinion regarding clinical practicalities in a structured interview, and average health care costs will be calculated.

Discussion

Principal Findings

Personalized lifestyle advice and promotion of self-management can help patients change their health behavior and improve glucose regulation. Digital tools have great potential in supporting patient self-management due to the effectiveness, low costs, 24/7 availability, and the option of dynamic automated feedback. However, reports documenting the impact of interventions incorporating multiple lifestyle modalities on glucose control are, to our knowledge, not available. Here, we developed an integrated, eHealth-supported, educational care pathway for people with T2D following the CeHReS roadmap and using a scoping review about DSME and eHealth, past experiences of eHealth practices in our hospital, focus groups with HCPs, and a patient panel. The care pathway aims to empower patients with T2D to self-manage their disease by providing them with direct feedback on their personal health behavior in relation to contemporaneous glucose levels.

HCPs and patients thought the concept of the Diabetes Box to be feasible, acceptable, and useful. The main strengths of the Diabetes Box were considered to be the integration of direct biofeedback on personal behavior, the focus on goal setting, and patient activation.

Comparison to Prior Work

The direct biofeedback regarding the impact of behavior on glucose concentrations was believed to be crucial to provide patients with insight into the relationships between their health behavior and glycemic control. A similar conclusion was drawn in an earlier study where patients with T2D were motivated to exercise while using CGM and accelerometer technology [64]. In many studies, data on lifestyle parameters were entered manually or via voice recording [22,24]. A Korean study showed input rates of diet and exercise of 24.9% and 5.3%, respectively [22]. Our study uses automatic input of steps, sleep, blood pressure, and glucose levels facilitating data gathering by participants. Diet was tracked using photographs. These photos were not analyzed for caloric content or carbohydrates but were used to provide insight into glucose-level fluctuations caused by certain food types. Beyond automated recording of behavioral and biological parameters, our app enables combining all lifestyle parameters with continuous glucose levels to create easily interpretable relations between lifestyle and glucose levels.

Regarding these lifestyle components, other interventions for people with T2D focus primarily on diet and exercise [23,24,65]. In one German study, stress management was included in the educational material, but stress or mood was not measured during the study [66]. The lifestyle components on which feedback should be provided include diet, physical activity, stress management, and sleep. Chronic stress may be a less obvious yet important disruptor of glucose control, as indicated by previous research [67,68]. Indeed, a recent meta-analysis revealed that stress reduction therapy improves glycemic control in people with diabetes [46]. The Diabetes Box gives direct biofeedback on diet, exercise, sleep, and stress. The effect of direct biofeedback on personal behavior is further enhanced by structured and tailored education. This is important, as physiological responses to lifestyle changes are often determined by personal characteristics [14,15]. Other studies often use one-size-fits-all education or even automated SMS text messages [24,65,69]. During the educational consultations in the Diabetes Box, HCPs can inform patients regarding the effects of their personal health behavior on metabolic control. The education in the Diabetes Box was designed to be simple, patient-centered, and multimodal, which is in line with the literature on successful patient education [70,71].

In addition, most of the existing tools are used in a research setting, and the challenge is to integrate these tools into regular medical care. A recent Dutch study showed that following a 2-year multicomponent lifestyle program outside of regular medical care could reduce medication use. In this setting, 71% of insulin users could stop insulin, and 28% of participants could stop glucose-lowering medication altogether. It must be stated though that only 234 of 438 starting participants were used in the final analyses [23]. When using these tools as regular medical care, all patients with T2D will follow the program instead of a selection of the more motivated patients. Our setup is to use this tool as regular medical care for all patients with T2D. The participatory development with an entire endocrinology team working in diabetes care can improve the chance of successful implementation. We are curious if similar results will be achieved when a multicomponent lifestyle program is integrated into regular medical care.

Studies have shown that people with higher levels of lifestyle-related knowledge (eg, influence of diet on glucose levels) tend to make healthier choices to improve their glycemic control [72]. However, better education and insight do not necessarily translate into behavior change [73]. To stimulate patients to change their behavior, goal setting and activation are integrated into all components of the Diabetes Box. All videos end with an assignment to self-monitor specific behavioral and biological parameters in preparation of the next consultation, and a separate video about goal setting is included. Furthermore, the individual consultations with the dietitian revolve around diet but also cover reflection on goals set by a patient. It is difficult to empower people with insufficient diabetes-related knowledge to manage their disease [74]. Therefore, we believe it is the combination of direct feedback, structured and tailored education, and targeted patient activation that grants the Diabetes Box its great potential.

The participatory development process played a critical role in the realization of the Diabetes Box. Involving all stakeholders from the start proved very fruitful, as it clearly facilitated the creation of a program that fits all stakeholder demands. The CeHReS roadmap was very helpful as well. It provided handholds and courses of action, which make it easier to make and measure progress. In addition, the value specification generated a concrete set of wishes from the key stakeholders that could be used to fall back and make decisions.

XSL•FO
RenderX

Limitations and Strengths

Obviously, there are at least 2 issues that limit the broad-based application of the Diabetes Box for the time being. First of all, the feasibility of the program as well as its impact on metabolic control and quality of life of patients with T2D needs to be evaluated in clinical practice. As the program was primarily developed by stakeholders employed by a third-line, academic medical institution, it also needs to be tested if it works for patients under regular surveillance by primary care (n>1 million, 90% of patients with T2D are treated by their general practitioners in the Netherlands). Second, although the LUMCCare app was created as a "white label" app, which means that it is relatively easy to adapt external characteristics, it was designed as part of the local (Leiden University Medical Center) infrastructure. Use by other institutions would therefore probably require modifications. We are willing to help and assist hospitals and other health care institutions that want to implement the Diabetes Box into their regular medical care for people with T2D. The challenges we foresee are training personnel and integrating the Diabetes Box into their daily workflow. The type of specialist who provides the consultations can be changed depending on what professionals are motivated and at hand. In addition, the content of the educational material can be altered to better fit the personal approach of the professional providing the education. On a technological basis, challenges also exist. The app is white label and can be easily adapted to accommodate the look and feel of other institutions. However, the data generated in the app have to be made available for the HCPs involved. This will most commonly involve integration into the electronic medical records, which is a process that costs both time and money. In the near future, the Diabetes Box will be tested in a single-center, mixed methods, sequential explanatory pilot study including approximately 32 patients with T2D, with the primary aim to assess its feasibility, acceptability, and usability. Secondary objectives will be to evaluate its impact on the "time in range" of glucose levels and perceived learning. Subsequently, in case of promising results, the Diabetes Box will be tested for efficacy in a larger, multicenter (including primary care) intervention study.

Conclusions

We have developed a unique care pathway in close collaboration with relevant stakeholders in order to ensure a good fit. The combined effects of direct biofeedback on personal behavior, structured and tailored education, and goal setting should empower people with T2D to improve their self-management and glycemic control. A pilot study is planned to assess feasibility, acceptability, and usability in more detail.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Description of behavior change techniques used in the Diabetes Box.
[DOCX File , 18 KB - humanfactors_v11i1e45055_app1.docx]

References

1. IDF diabetes atlas. International Diabetes Federation. 2019. URL: https://diabetesatlas.org/idfawp/resource-files/2019/07/IDF_diabetes_atlas_ninth_edition_en.pdf [accessed 2023-04-12]
2. Diabetes fact sheet. World Health Organization. 2020. URL: https://www.who.int/news-room/fact-sheets/detail/diabetes [accessed 2023-03-11]
3. Ford ES, Bergmann MM, Kröger J, Schienkiewitz A, Weikert C, Boeing H. Healthy living is the best revenge: findings from the European Prospective Investigation Into Cancer and Nutrition-Potsdam study. Arch Intern Med 2009;169(15):1355-1362. [doi: 10.1001/archinternmed.2009.237] [Medline: 19667296]
4. Uusitupa M, Khan TA, Viguiliouk E, Kahleova H, Rivellese AA, Hermansen K, et al. Prevention of type 2 diabetes by lifestyle changes: a systematic review and meta-analysis. Nutrients 2019;11(11):2611 [FREE Full text] [doi: 10.3390/nu11112611] [Medline: 31683759]
5. Churuangsuk C, Hall J, Reynolds A, Griffin SJ, Combet E, Lean MEJ. Diets for weight management in adults with type 2 diabetes: an umbrella review of published meta-analyses and systematic review of trials of diets for diabetes remission. Diabetologia 2022;65(1):14-36 [FREE Full text] [doi: 10.1007/s00125-021-05577-2] [Medline: 34796367]
6. da Rocha RB, Silva CS, Cardoso VS. Self-care in adults with type 2 diabetes mellitus: a systematic review. Curr Diabetes Rev 2020;16(6):598-607. [doi: 10.2174/1573399815666190702161849] [Medline: 31267873]
7. Jarvie JL, Pandey A, Ayers CR, McGavock JM, Sénéchal M, Berry JD, et al. Aerobic fitness and adherence to guideline-recommended minimum physical activity among ambulatory patients with type 2 diabetes mellitus. Diabetes Care 2019;42(7):1333-1339 [FREE Full text] [doi: 10.2337/dc18-2634] [Medline: 31221698]
8. Rivellese AA, Boemi M, Cavalot F, Costagliola L, De Feo P, Miccoli R, et al. Dietary habits in type II diabetes mellitus: how is adherence to dietary recommendations? Eur J Clin Nutr 2008;62(5):660-664. [doi: 10.1038/sj.ejcn.1602755] [Medline: 17426738]

XSL•FO
RenderX

9. Houle J, Beaulieu MD, Chiasson J, Lespérance F, Côté J, Strychar I, et al. Glycaemic control and self-management behaviours in type 2 diabetes: results from a 1-year longitudinal cohort study. Diabet Med 2015;32(9):1247-1254. [doi: 10.1111/dme.12686] [Medline: 25581545]
10. Alonso-Fernández M, Mancera-Romero J, Mediavilla-Bravo JJ, Comas-Samper JM, López-Simarro F, Pérez-Unanua MP, et al. Glycemic control and use of A1c in primary care patients with type 2 diabetes mellitus. Prim Care Diabetes 2015;9(5):385-391. [doi: 10.1016/j.pcd.2015.01.006] [Medline: 25686480]
11. American Diabetes Association. 6. Glycemic targets: standards of medical care in diabetes-2021. Diabetes Care 2021;44(Suppl 1):S73-S84 [FREE Full text] [doi: 10.2337/dc21-S006] [Medline: 33298417]
12. Deakin T, McShane CE, Cade JE, Williams RDRR. Group based training for self-management strategies in people with type 2 diabetes mellitus. Cochrane Database Syst Rev 2005(2):CD003417. [doi: 10.1002/14651858.CD003417.pub2] [Medline: 15846663]
13. Almutairi N, Hosseinzadeh H, Gopaldasani V. The effectiveness of patient activation intervention on type 2 diabetes mellitus glycemic control and self-management behaviors: a systematic review of RCTs. Prim Care Diabetes 2020;14(1):12-20. [doi: 10.1016/j.pcd.2019.08.009] [Medline: 31543458]
14. Berry SE, Valdes AM, Drew DA, Asnicar F, Mazidi M, Wolf J, et al. Human postprandial responses to food and potential for precision nutrition. Nat Med 2020;26(6):964-973 [FREE Full text] [doi: 10.1038/s41591-020-0934-0] [Medline: 32528151]
15. Zeevi D, Korem T, Zmora N, Israeli D, Rothschild D, Weinberger A, et al. Personalized nutrition by prediction of glycemic responses. Cell 2015;163(5):1079-1094 [FREE Full text] [doi: 10.1016/j.cell.2015.11.001] [Medline: 26590418]
16. Klein HA, Lippa KD. Type 2 diabetes self-management: controlling a dynamic system. J Cogn Eng Decis Mak 2008;2(1):48-62. [doi: 10.1518/155534308x284363]
17. Gershkowitz BD, Hillert CJ, Crotty BH. Digital coaching strategies to facilitate behavioral change in type 2 diabetes: a systematic review. J Clin Endocrinol Metab 2021;106(4):e1513-e1520 [FREE Full text] [doi: 10.1210/clinem/dgaa850] [Medline: 33206975]
18. Park C, Le QA. The effectiveness of continuous glucose monitoring in patients with type 2 diabetes: a systematic review of literature and meta-analysis. Diabetes Technol Ther 2018;20(9):613-621. [doi: 10.1089/dia.2018.0177] [Medline: 30095980]
19. Jaana M, Paré G. Home telemonitoring of patients with diabetes: a systematic assessment of observed effects. J Eval Clin Pract 2007;13(2):242-253. [doi: 10.1111/j.1365-2753.2006.00686.x] [Medline: 17378871]
20. Wu X, Guo X, Zhang Z. The efficacy of mobile phone apps for lifestyle modification in diabetes: systematic review and meta-analysis. JMIR Mhealth Uhealth 2019;7(1):e12297 [FREE Full text] [doi: 10.2196/12297] [Medline: 30664494]
21. Hou C, Carter B, Hewitt J, Francisa T, Mayor S. Do mobile phone applications improve glycemic control (HbA1c) in the self-management of diabetes? A systematic review, meta-analysis, and GRADE of 14 randomized trials. Diabetes Care 2016;39(11):2089-2095 [FREE Full text] [doi: 10.2337/dc16-0346] [Medline: 27926892]
22. Oh SW, Kim KK, Kim SS, Park SK, Park S. Effect of an integrative mobile health intervention in patients with hypertension and diabetes: crossover study. JMIR Mhealth Uhealth 2022;10(1):e27192 [FREE Full text] [doi: 10.2196/27192] [Medline: 35014961]
23. Pot GK, Battjes-Fries MC, Patijn ON, van der Zijl N, Pijl H, Voshol P. Lifestyle medicine for type 2 diabetes: practice-based evidence for long-term efficacy of a multicomponent lifestyle intervention (Reverse Diabetes2 Now). BMJ Nutr Prev Health 2020;3(2):188-195 [FREE Full text] [doi: 10.1136/bmjnph-2020-000081] [Medline: 33521528]
24. Lee EY, Cha SA, Yun JS, Lim SY, Lee JH, Ahn YB, et al. Efficacy of personalized diabetes self-care using an electronic medical record-integrated mobile app in patients with type 2 diabetes: 6-month randomized controlled trial. J Med Internet Res 2022;24(7):e37430 [FREE Full text] [doi: 10.2196/37430] [Medline: 35900817]
25. Ruissen MM, Rodriguez-Gutierrez R, Montori VM, Kunneman M. Making diabetes care fit—are we making progress? Front Clin Diabetes Healthc 2021;2:658817 [FREE Full text] [doi: 10.3389/fcdhc.2021.658817] [Medline: 36994329]
26. Eysenbach G. Medicine 2.0: social networking, collaboration, participation, apomediation, and openness. J Med Internet Res 2008;10(3):e22 [FREE Full text] [doi: 10.2196/jmir.1030] [Medline: 18725354]
27. Granja C, Janssen W, Johansen MA. Factors determining the success and failure of eHealth interventions: systematic review of the literature. J Med Internet Res 2018;20(5):e10235 [FREE Full text] [doi: 10.2196/10235] [Medline: 29716883]
28. Wachter RM. Making IT work: harnessing the power of health information technology to improve care in England. National Advisory Group on Health Information Technology in England. 2016. URL: https://assets.publishing.service.gov.uk/media/5a8091afe5274a2e87dba8f2/Wachter_Review_Accessible.pdf [accessed 2023-01-26]
29. van Velsen L, Evers M, Bara C, Op den Akker H, Boerema S, Hermens H. Understanding the acceptance of an eHealth technology in the early stages of development: an end-user walkthrough approach and two case studies. JMIR Form Res 2018;2(1):e10474 [FREE Full text] [doi: 10.2196/10474] [Medline: 30684434]
30. Hennemann S, Beutel ME, Zwerenz R. Ready for eHealth? Health professionals' acceptance and adoption of eHealth interventions in inpatient routine care. J Health Commun 2017;22(3):274-284. [doi: 10.1080/10810730.2017.1284286] [Medline: 28248626]

XSL•FO
RenderX

31. van Gemert-Pijnen JEWC, Nijland N, van Limburg M, Ossebaard HC, Kelders SM, Eysenbach G, et al. A holistic framework to improve the uptake and impact of eHealth technologies. J Med Internet Res 2011 05;13(4):e111 [FREE Full text] [doi: 10.2196/jmir.1672] [Medline: 22155738]
32. Westphaln KK, Regoeczi W, Masotya M, Vazquez-Westphaln B, Lounsbury K, McDavid L, et al. From Arksey and O'Malley and beyond: customizations to enhance a team-based, mixed approach to scoping review methodology. MethodsX 2021;8:101375 [FREE Full text] [doi: 10.1016/j.mex.2021.101375] [Medline: 34430271]
33. Treskes RW, van den Akker-van Marle ME, van Winden L, van Keulen N, van der Velde ET, Beeres S, et al. The Box—eHealth in the outpatient clinic follow-up of patients with acute myocardial infarction: cost-utility analysis. J Med Internet Res 2022;24(4):e30236 [FREE Full text] [doi: 10.2196/30236] [Medline: 35468091]
34. Biersteker TE, Boogers MJ, Schalij MJ, Penning de Vries B, Groenwold RHH, van Alem AP, et al. Mobile health vs. standard care after cardiac surgery: results of The Box 2.0 study. Europace 2023;25(1):49-58 [FREE Full text] [doi: 10.1093/europace/euac115] [Medline: 35951658]
35. Dirikgil E, Roos R, Groeneveld GH, Heringhaus C, Silven AV, Petrus AHJ, et al. Home monitoring reduced short stay admissions in suspected COVID-19 patients: COVID-box project. Eur Respir J 2021;58(2):2100636 [FREE Full text] [doi: 10.1183/13993003.00636-2021] [Medline: 33795321]
36. Stuckey HL, Vallis M, Burns KK, Mullan-Jensen CB, Reading JM, Kalra S, et al. "I do my best to listen to patients": qualitative insights into DAWN2 (diabetes psychosocial care from the perspective of health care professionals in the second diabetes attitudes, wishes and needs study). Clin Ther 2015;37(9):1986-1998.e12 [FREE Full text] [doi: 10.1016/j.clinthera.2015.06.010] [Medline: 26169765]
37. Raaijmakers LG, Hamers FJ, Martens MK, Bagchus C, de Vries NK, Kremers SP. Perceived facilitators and barriers in diabetes care: a qualitative study among health care professionals in the Netherlands. BMC Fam Pract 2013;14:114 [FREE Full text] [doi: 10.1186/1471-2296-14-114] [Medline: 23937325]
38. Yoon S, Ng JH, Kwan YH, Low LL. Healthcare professionals' views of factors influencing diabetes self-management and the utility of a mHealth application and its features to support self-care. Front Endocrinol (Lausanne) 2022;13:793473 [FREE Full text] [doi: 10.3389/fendo.2022.793473] [Medline: 35282452]
39. Sun Y, You W, Almeida F, Estabrooks P, Davy B. The effectiveness and cost of lifestyle interventions including nutrition education for diabetes prevention: a systematic review and meta-analysis. J Acad Nutr Diet 2017;117(3):404-421.e36 [FREE Full text] [doi: 10.1016/j.jand.2016.11.016] [Medline: 28236962]
40. García-Molina L, Lewis-Mikhael AM, Riquelme-Gallego B, Cano-Ibáñez N, Oliveras-López MJ, Bueno-Cavanillas A. Improving type 2 diabetes mellitus glycaemic control through lifestyle modification implementing diet intervention: a systematic review and meta-analysis. Eur J Nutr 2020;59(4):1313-1328. [doi: 10.1007/s00394-019-02147-6] [Medline: 31781857]
41. Hemmingsen B, Gimenez-Perez G, Mauricio D, Figuls MRI, Metzendorf MI, Richter B. Diet, physical activity or both for prevention or delay of type 2 diabetes mellitus and its associated complications in people at increased risk of developing type 2 diabetes mellitus. Cochrane Database Syst Rev 2017;12(12):CD003054 [FREE Full text] [doi: 10.1002/14651858.CD003054.pub4] [Medline: 29205264]
42. Kumar AS, Maiya AG, Shastry BA, Vaishali K, Ravishankar N, Hazari A, et al. Exercise and insulin resistance in type 2 diabetes mellitus: a systematic review and meta-analysis. Ann Phys Rehabil Med 2019;62(2):98-103 [FREE Full text] [doi: 10.1016/j.rehab.2018.11.001] [Medline: 30553010]
43. Esposito K, Maiorino MI, Bellastella G, Chiodini P, Panagiotakos D, Giugliano D. A journey into a Mediterranean diet and type 2 diabetes: a systematic review with meta-analyses. BMJ Open 2015;5(8):e008222 [FREE Full text] [doi: 10.1136/bmjopen-2015-008222] [Medline: 26260349]
44. Cappuccio FP, D'Elia L, Strazzullo P, Miller MA. Quantity and quality of sleep and incidence of type 2 diabetes: a systematic review and meta-analysis. Diabetes Care 2010;33(2):414-420 [FREE Full text] [doi: 10.2337/dc09-1124] [Medline: 19910503]
45. Uchendu C, Blake H. Effectiveness of cognitive-behavioural therapy on glycaemic control and psychological outcomes in adults with diabetes mellitus: a systematic review and meta-analysis of randomized controlled trials. Diabet Med 2017;34(3):328-339. [doi: 10.1111/dme.13195] [Medline: 27472405]
46. Ni Y, Ma L, Li J. Effects of mindfulness-based stress reduction and mindfulness-based cognitive therapy in people with diabetes: a systematic review and meta-analysis. J Nurs Scholarsh 2020;52(4):379-388. [doi: 10.1111/jnu.12560] [Medline: 32406186]
47. Seto E, Istepanian RSH, Cafazzo JA, Logan A, Sungoor A. UK and Canadian perspectives of the effectiveness of mobile diabetes management systems. Annu Int Conf IEEE Eng Med Biol Soc 2009;2009:6584-6587. [doi: 10.1109/IEMBS.2009.5333998] [Medline: 19964700]
48. Janapala RN, Jayaraj JS, Fathima N, Kashif T, Usman N, Dasari A, et al. Continuous glucose monitoring versus self-monitoring of blood glucose in type 2 diabetes mellitus: a systematic review with meta-analysis. Cureus 2019;11(9):e5634 [FREE Full text] [doi: 10.7759/cureus.5634] [Medline: 31700737]

XSL•FO
RenderX

49. Beck RW, Riddlesworth TD, Ruedy K, Ahmann A, Haller S, Kruger D, et al. Continuous glucose monitoring versus usual care in patients with type 2 diabetes receiving multiple daily insulin injections: a randomized trial. Ann Intern Med 2017;167(6):365-374. [doi: 10.7326/M16-2855] [Medline: 28828487]
50. Chomutare T, Fernandez-Luque L, Arsand E, Hartvigsen G. Features of mobile diabetes applications: review of the literature and analysis of current applications compared against evidence-based guidelines. J Med Internet Res 2011;13(3):e65 [FREE Full text] [doi: 10.2196/jmir.1874] [Medline: 21979293]
51. Hall KS, Hyde ET, Bassett DR, Carlson SA, Carnethon MR, Ekelund U, et al. Systematic review of the prospective association of daily step counts with risk of mortality, cardiovascular disease, and dysglycemia. Int J Behav Nutr Phys Act 2020;17(1):78 [FREE Full text] [doi: 10.1186/s12966-020-00978-9] [Medline: 32563261]
52. Kraus WE, Janz KF, Powell KE, Campbell WW, Jakicic JM, Troiano RP, et al. Daily step counts for measuring physical activity exposure and its relation to health. Med Sci Sports Exerc 2019;51(6):1206-1212 [FREE Full text] [doi: 10.1249/MSS.0000000000001932] [Medline: 31095077]
53. Tucker KL, Sheppard JP, Stevens R, Bosworth HB, Bove A, Bray EP, et al. Self-monitoring of blood pressure in hypertension: a systematic review and individual patient data meta-analysis. PLoS Med 2017;14(9):e1002389 [FREE Full text] [doi: 10.1371/journal.pmed.1002389] [Medline: 28926573]
54. Logan AG, Irvine MJ, McIsaac WJ, Tisler A, Rossos PG, Easty A, et al. Effect of home blood pressure telemonitoring with self-care support on uncontrolled systolic hypertension in diabetics. Hypertension 2012;60(1):51-517 [FREE Full text] [doi: 10.1161/HYPERTENSIONAHA.111.188409] [Medline: 22615116]
55. Puhl RM, Himmelstein MS, Hateley-Browne JL, Speight J. Weight stigma and diabetes stigma in U.S. adults with type 2 diabetes: associations with diabetes self-care behaviors and perceptions of health care. Diabetes Res Clin Pract 2020;168:108387. [doi: 10.1016/j.diabres.2020.108387] [Medline: 32858100]
56. Hou C, Xu Q, Diao S, Hewitt J, Li J, Carter B. Mobile phone applications and self-management of diabetes: a systematic review with meta-analysis, meta-regression of 21 randomized trials and GRADE. Diabetes Obes Metab 2018;20(8):2009-2013. [doi: 10.1111/dom.13307] [Medline: 29582538]
57. Torbjørnsen A, Ribu L, Rønnevig M, Grøttland A, Helseth S. Users' acceptability of a mobile application for persons with type 2 diabetes: a qualitative study. BMC Health Serv Res 2019;19(1):641 [FREE Full text] [doi: 10.1186/s12913-019-4486-2] [Medline: 31492176]
58. Jeffrey B, Bagala M, Creighton A, Leavey T, Nicholls S, Wood C, et al. Mobile phone applications and their use in the self-management of Type 2 Diabetes Mellitus: a qualitative study among app users and non-app users. Diabetol Metab Syndr 2019;11:84 [FREE Full text] [doi: 10.1186/s13098-019-0480-4] [Medline: 31636719]
59. Rollo ME, Ash S, Lyons-Wall P, Russell AW. Evaluation of a mobile phone image-based dietary assessment method in adults with type 2 diabetes. Nutrients 2015;7(6):4897-4910 [FREE Full text] [doi: 10.3390/nu7064897] [Medline: 26091234]
60. Thompson FE, Subar AF. Chapter 1—Dietary assessment methodology. In: Coulston AM, Boushey CJ, Ferruzzi MG, Delahanty LM, editors. Nutrition in the Prevention and Treatment of Disease. 4th Edition. Cambridge, MA: Academic Press; 2017:5-48.
61. Schap TE, Six BL, Delp EJ, Ebert DS, Kerr DA, Boushey CJ. Adolescents in the United States can identify familiar foods at the time of consumption and when prompted with an image 14 h postprandial, but poorly estimate portions. Public Health Nutr 2011;14(7):1184-1191 [FREE Full text] [doi: 10.1017/S1368980010003794] [Medline: 21324224]
62. Treskes RW, van Winden LAM, van Keulen N, van der Velde ET, Beeres SLMA, Atsma DE, et al. Effect of smartphone-enabled health monitoring devices vs regular follow-up on blood pressure control among patients after myocardial infarction: a randomized clinical trial. JAMA Netw Open 2020;3(4):e202165 [FREE Full text] [doi: 10.1001/jamanetworkopen.2020.2165] [Medline: 32297946]
63. Michie S, Ashford SA, Sniehotta FF, Dombrowski SU, Bishop A, French DP. A refined taxonomy of behaviour change techniques to help people change their physical activity and healthy eating behaviours: the CALO-RE taxonomy. Psychol Health 2011;26(11):1479-1498. [doi: 10.1080/08870446.2010.540664] [Medline: 21678185]
64. Allen NA, Jacelon CS, Chipkin SR. Feasibility and acceptability of continuous glucose monitoring and accelerometer technology in exercising individuals with type 2 diabetes. J Clin Nurs 2009;18(3):373-383 [FREE Full text] [doi: 10.1111/j.1365-2702.2008.02533.x] [Medline: 19191984]
65. Chen L, Chuang LM, Chang CH, Wang CS, Wang IC, Chung Y, et al. Evaluating self-management behaviors of diabetic patients in a telehealthcare program: longitudinal study over 18 months. J Med Internet Res 2013;15(12):e266 [FREE Full text] [doi: 10.2196/jmir.2699] [Medline: 24323283]
66. von Storch K, Graaf E, Wunderlich M, Rietz C, Polidori MC, Woopen C. Telemedicine-assisted self-management program for type 2 diabetes patients. Diabetes Technol Ther 2019;21(9):514-521. [doi: 10.1089/dia.2019.0056] [Medline: 31287736]
67. Derosa G, Salvadeo SAT, Mereu R, D'Angelo A, Ciccarelli L, Piccinni MN, et al. Continuous glucose monitoring system in free-living healthy subjects: results from a pilot study. Diabetes Technol Ther 2009;11(3):159-169. [doi: 10.1089/dia.2008.0101] [Medline: 19267581]
68. Sevil M, Rashid M, Hajizadeh I, Park M, Quinn L, Cinar A. Physical activity and psychological stress detection and assessment of their effects on glucose concentration predictions in diabetes management. IEEE Trans Biomed Eng 2021;68(7):2251-2260 [FREE Full text] [doi: 10.1109/TBME.2020.3049109] [Medline: 33400644]

XSL•FO
RenderX

69. Bretschneider MP, Klásek J, Karbanová M, Timpel P, Herrmann S, Schwarz PEH. Impact of a digital lifestyle intervention on diabetes self-management: a pilot study. Nutrients 2022;14(9):1810 [FREE Full text] [doi: 10.3390/nu14091810] [Medline: 35565782]
70. Cutilli CC. Excellence in patient education: evidence-based education that "sticks" and improves patient outcomes. Nurs Clin North Am 2020;55(2):267-282. [doi: 10.1016/j.cnur.2020.02.007] [Medline: 32389259]
71. Friedman AJ, Cosby R, Boyko S, Hatton-Bauer J, Turnbull G. Effective teaching strategies and methods of delivery for patient education: a systematic review and practice guideline recommendations. J Cancer Educ 2011;26(1):12-21. [doi: 10.1007/s13187-010-0183-x] [Medline: 21161465]
72. Cheng LJ, Wang W, Lim ST, Wu VX. Factors associated with Glycaemic control in patients with diabetes mellitus: a systematic literature review. J Clin Nurs 2019;28(9-10):1433-1450. [doi: 10.1111/jocn.14795] [Medline: 30667583]
73. Kelly MP, Barker M. Why is changing health-related behaviour so difficult? Public Health 2016;136:109-116 [FREE Full text] [doi: 10.1016/j.puhe.2016.03.030] [Medline: 27184821]
74. Lambrinou E, Hansen TB, Beulens JW. Lifestyle factors, self-management and patient empowerment in diabetes care. Eur J Prev Cardiol 2019;26(2_suppl):55-63 [FREE Full text] [doi: 10.1177/2047487319885455] [Medline: 31766913]

Abbreviations

CeHReS: Center for eHealth and Wellbeing Research
CGM: continuous glucose monitoring
DSME: diabetes self-management education
HCP: health care professional
T2D: type 2 diabetes

Edited by A Kushniruk; submitted 15.12.22; peer-reviewed by V Buss, SM Ching , SLS Drop; comments to author 14.05.23; revised version received 17.09.23; accepted 20.11.23; published 31.05.24.

Please cite as:
de Frel DL, Schroijen MA, Aardoom JJ, van Gils W, Huisman SD, Janssen VR, Versluis A, Kleinsmann MS, Atsma DE, Pijl H Participatory Development of an Integrated, eHealth-Supported, Educational Care Pathway (Diabetes Box) for People With Type 2 Diabetes: Development and Usability Study
JMIR Hum Factors 2024;11:e45055
URL: https://humanfactors.jmir.org/2024/1/e45055
doi:10.2196/45055
PMID:38819880

©Daan Leonhard de Frel, Mariëlle A Schroijen, Jiska J Aardoom, Wesley van Gils, Sasja D Huisman, Veronica R Janssen, Anke Versluis, Maaike S Kleinsmann, Douwe E Atsma, Hanno Pijl. Originally published in JMIR Human Factors (https://humanfactors.jmir.org), 31.05.2024. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on https://humanfactors.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Assessing the Utility, Impact, and Adoption Challenges of an Artificial Intelligence–Enabled Prescription Advisory Tool for Type 2 Diabetes Management: Qualitative Study

Sungwon Yoon[1,2], MPH, PhD; Hendra Goh[1], MPH; Phong Ching Lee[3], MBChB, MRCP; Hong Chang Tan[3], MBBS, MRCP; Ming Ming Teh[3], MB BCh BAO, MRCP, MD; Dawn Shao Ting Lim[3], MBBS, MRCP, MMed; Ann Kwee[3], MB BCh BAO, MRCP, MMed; Chandran Suresh[3], MBBS, MRCP; David Carmody[3], MB BCh BAO, LRCP & SI, MRCP, MD; Du Soon Swee[3], MBBS, MRCP; Sarah Ying Tse Tan[3], MBBS, MRCP, MMed; Andy Jun-Wei Wong[3], MBBS; Charlotte Hui-Min Choo[3], MBBS; Zongwen Wee[3], MBBS, MRCP, MMed; Yong Mong Bee[3], MBBS, MRCP

[1]Health Services and Systems Research, Duke-NUS Medical School, Singapore, Singapore

[2]Centre for Population Health Research and Implementation, SingHealth Regional Health System, SingHealth, Singapore, Singapore

[3]Department of Endocrinology, Singapore General Hospital, Singapore, Singapore

Corresponding Author:
Sungwon Yoon, MPH, PhD
Health Services and Systems Research
Duke-NUS Medical School
8 College Road
Singapore, 169857
Singapore
Phone: 65 66013198
Email: sungwon.yoon@duke-nus.edu.sg

Abstract

Background: The clinical management of type 2 diabetes mellitus (T2DM) presents a significant challenge due to the constantly evolving clinical practice guidelines and growing array of drug classes available. Evidence suggests that artificial intelligence (AI)–enabled clinical decision support systems (CDSSs) have proven to be effective in assisting clinicians with informed decision-making. Despite the merits of AI-driven CDSSs, a significant research gap exists concerning the early-stage implementation and adoption of AI-enabled CDSSs in T2DM management.

Objective: This study aimed to explore the perspectives of clinicians on the use and impact of the AI-enabled Prescription Advisory (APA) tool, developed using a multi-institution diabetes registry and implemented in specialist endocrinology clinics, and the challenges to its adoption and application.

Methods: We conducted focus group discussions using a semistructured interview guide with purposively selected endocrinologists from a tertiary hospital. The focus group discussions were audio-recorded and transcribed verbatim. Data were thematically analyzed.

Results: A total of 13 clinicians participated in 4 focus group discussions. Our findings suggest that the APA tool offered several useful features to assist clinicians in effectively managing T2DM. Specifically, clinicians viewed the AI-generated medication alterations as a good knowledge resource in supporting the clinician's decision-making on drug modifications at the point of care, particularly for patients with comorbidities. The complication risk prediction was seen as positively impacting patient care by facilitating early doctor-patient communication and initiating prompt clinical responses. However, the interpretability of the risk scores, concerns about overreliance and automation bias, and issues surrounding accountability and liability hindered the adoption of the APA tool in clinical practice.

Conclusions: Although the APA tool holds great potential as a valuable resource for improving patient care, further efforts are required to address clinicians' concerns and improve the tool's acceptance and applicability in relevant contexts.

(JMIR Hum Factors 2024;11:e50939) doi:10.2196/50939

XSL•FO
RenderX

KEYWORDS

clinical decision support system; artificial intelligence; endocrinology; diabetes management; human factors

Introduction

Diabetes mellitus is a chronic condition that affects millions of people worldwide. In Singapore, the prevalence of diabetes is estimated to surpass 400,000, with 1 out of 3 individuals at risk of developing the condition [1]. Uncontrolled diabetes can lead to various complications, such as neuropathy, retinopathy, and nephropathy. Diabetes is primarily associated with cardiovascular diseases, particularly ischemic heart disease and myocardial infarction, which account for most of the mortality cases in patients with diabetes [2,3].

Managing diabetes clinically poses a considerable challenge due to its complex nature. The treatment of diabetes involves achieving specific targets, such as optimal control of glycemia, blood pressure, and lipid levels, primarily relying on laboratory tests [4]. Regular review of test results and subsequent treatment adjustments are important in minimizing the risk of long-term complications and aligning with recommended targets [5]. However, the need to monitor multiple laboratory markers during clinical consultations can impose a cognitive burden. Furthermore, incomplete integration of critical patient data into the electronic medical record (EMR) can lead to errors in disease monitoring, compromising the quality of patient care [6].

Evidence suggests that clinical decision support systems (CDSSs) can assist clinicians in effectively monitoring patient data and making accurate and informed treatment decisions [7-9]. Traditionally, CDSSs have relied on medical expertise and clinical practice guidelines. However, keeping CDSS content and knowledge up-to-date is increasingly challenging due to the evolving nature of clinical practices [10]. The advent of big data and machine learning has enabled the development of artificial intelligence (AI)-powered CDSSs, capable of diagnosing conditions, suggesting evidence-based treatment options and aiding in care planning [11,12]. Research shows that AI-powered CDSSs have improved the quality of diabetes care and patient outcomes [12,13].

Despite the positive impacts of AI-driven CDSSs on health care, fewer studies have examined human factors. In addition, several critical issues surrounding AI-powered clinical tools have been brought to attention, including concerns regarding the transparency of underlying algorithms, accountability, data privacy, and limited trust and applicability [10,14]. Although these studies provide essential insights into implementing AI-based CDSSs, a significant gap exists in research concerning the early-stage implementation of an AI-enabled CDSS specifically for type 2 diabetes mellitus (T2DM). Evaluating a CDSS in the early stages of implementation is of utmost importance to optimize its benefits and mitigate potential drawbacks, as it offers vital information on use, acceptability, and the challenges pertaining to human factors in real-world clinical settings.

To support clinicians in making better treatment decisions in T2DM management, the AI-enabled Prescription Advisory (APA) tool was developed and integrated within the endocrinology specialist clinics at the Diabetes & Metabolism Centre in Singapore General Hospital. To ensure that the tool is capable of scaling up and meeting the needs of its users, it is crucial to assess its appropriateness within the clinical context. Therefore, this study aims to explore the perspectives of clinicians regarding the use and impact of the APA tool while also identifying potential challenges associated with its adoption and application.

Methods

Overview

This study adopted qualitative research methodology to assess the usability of the tool. For rigor and transparency, we anchored our study according to the COREQ (Consolidated Criteria for Reporting Qualitative Research) checklist [15].

Development of an AI-Enabled Diabetes CDSS

The diabetes CDSS, also known as the APA tool, was developed using data gathered from the Singapore Health Services Diabetes Registry that comprised a total of 189,520 patients with diabetes. This data set included 6,407,958 outpatient visits spanning over 5 years from 2013 to 2018 [16]. For model development, 80% of the data set was used to build therapeutic recommendations, while the remaining 20% was used to test and validate the trained models. Three distinct therapeutic recommendation models were formulated for antiglycemic, antihypertensive, and lipid-lowering treatments. These models were created by integrating both a knowledge-driven approach and a data-driven approach. The knowledge-driven approach, initially drawing inputs from clinical guidelines and expert opinions, was used to identify potential therapeutic options. Subsequently, the data-driven approach that used deep learning techniques was used to select the identified therapeutic options based on anticipated clinical outcomes. To assess the performance of model's prediction, short-term outcomes compared therapeutic options between treatments that aligned with the model's recommendations and those that did not. Confounding factors were also accounted along the way and adjusted by stratification and multivariate regression. For evaluation of long-term outcomes, the rates of model-concordant treatments were computed by multivariate logistic regression to determine whether the combined treatments exhibited a positive impact on reducing the occurrence of long-term complications and mortality.

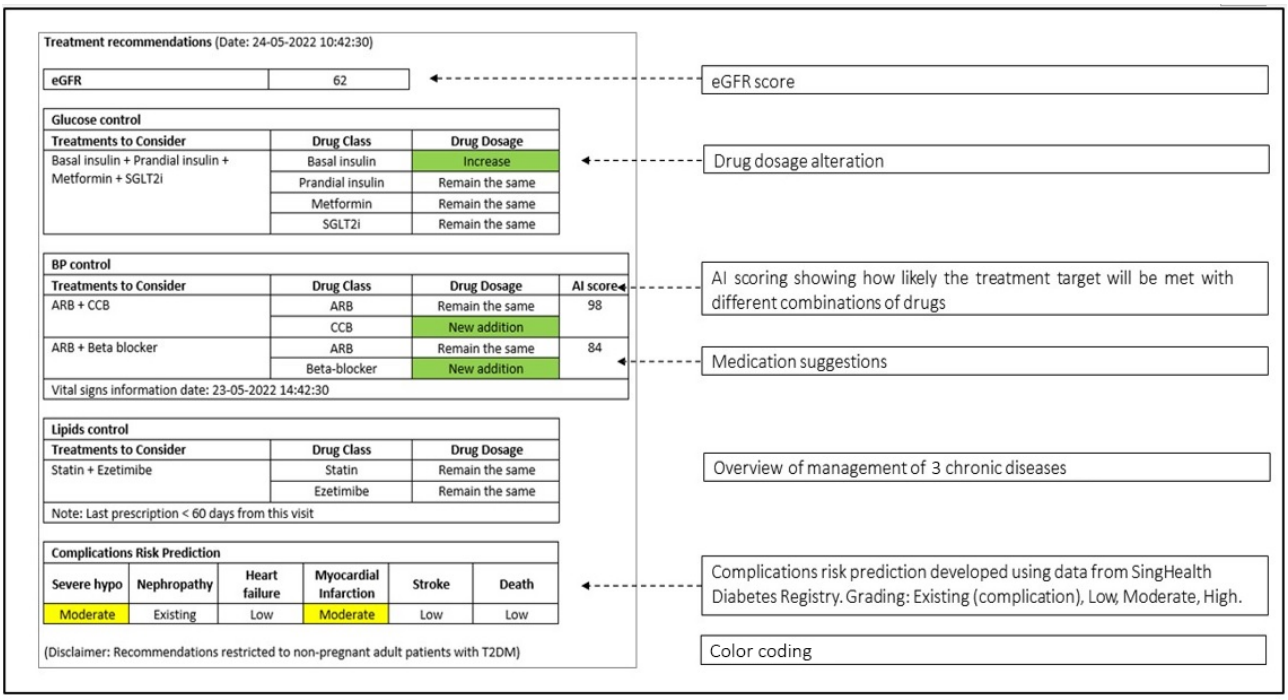
Features of the AI-Enabled Diabetes CDSS

Presently, the APA tool has been integrated into the EMR system at the Singapore General Hospital to provide tailored treatment recommendations for achieving target glucose, blood pressure, and cholesterol levels. It features 3 distinct AI components designed to improve patient outcomes and support clinical decision-making. The first component recommends drug classes based on laboratory markers such as glycated hemoglobin (HbA_{1c}), low-density lipoprotein cholesterol, and blood pressure measurements, aiding clinicians in selecting the

XSL•FO
RenderX

most appropriate drug classes to achieve glycemic, low-density lipoprotein cholesterol and blood pressure treatment targets. The second component generates an AI score that indicates the likelihood of reaching treatment targets when adopting a suggested new drug therapy. The third component generates AI-based diabetic complications risk predictions, providing specific complication risks associated with suboptimal diabetes treatment. All outputs generated by the AI model are color-coded for enhanced visibility. Figure 1 provides a snapshot of the APA tool's outputs, illustrating how clinical results extracted from a patient's EMR are presented.

Figure 1. Outputs of APA tool. The APA tool features include laboratory markers related to diabetes care, medication prescribing recommendations, color-coding to highlight changes, and diabetic complications risk prediction. AI: artificial intelligence; APA: AI-enabled Prescription Advisory; ARB: angiotensin receptor blockers; BP: blood pressure; CCB: calcium channel blockers; eGFR: estimated glomerular filtration rate; T2DM: type 2 diabetes mellitus.



***Note**: The APA tool features include laboratory markers related to diabetes care, medication prescribing recommendations, color-coding to highlight changes and diabetic complications risk prediction.*

Participants

Eligible participants were (1) clinicians trained in endocrinology, (2) currently employed full-time by the institution, (3) completed training in APA, and (4) used the APA for a minimum duration of 4 weeks. We purposively selected participants according to age, gender, and seniority level in the workplace to gain a range of perspectives. Participants were approached via email, and informed consent was obtained prior to enrollment into the study.

Study Procedure

Prospective participants were purposefully selected and approached by the research team to ensure their engagement in the study. Consented participants took part in a comprehensive group training session, encompassing the following key components: (1) a comprehensive overview of the tool's development and validation process, (2) exploration of the specific features and underlying knowledge rules that inform the therapeutic recommendations, (3) familiarization with the tool's output (ie, therapeutic recommendations), and (4) an interactive question and answer segment addressing general inquiries. The entirety of the training session was conducted over the course of 1 hour. Following the training, the participants were granted immediate access to the APA tool during their clinical consultations. It was emphasized that the participants had the freedom to decide on the final treatments for their patients with T2DM, regardless of the clinical recommendations provided by the APA tool. Each participant was given a minimum of 4 weeks of exposure to the tool in the clinic setting before his or her involvement in this study. Following the 4-week period, the participants were invited to partake in a focus group discussion (FGD).

Data Collection

A semistructured FGD guide was developed by drawing upon relevant literature and leveraging the expert knowledge of the study team [17,18]. FGD was chosen to foster a dynamic and interactive environment that encouraged the exploration of shared experiences and perspectives within the professional context, thereby providing a comprehensive understanding of the collective viewpoints users had on the APA tool. To attain a variety of opinions, the participants were recruited according to their seniority (registrars, associate consultants, consultants, and senior consultants). FGDs were conducted exclusively among participants holding equivalent hierarchical positions within the workplace, with the specific intention of minimizing the potential for power differentials and fostering open and

XSL•FO
RenderX

candid dialogue. Key topics of interest included (1) participants' firsthand experiences while using the APA tool, (2) perceptions and evaluations of the various features, (3) impacts of the APA tool on clinical practice, and (4) challenges related to the adoption and application. The interview guide underwent pilot-testing multiple iterations. Consented individuals were then invited to participate in virtual FGD (2-6 participants per session according to seniority) over Zoom (Zoom Video Communications, Inc) by a facilitator (HG) trained in qualitative research methodology. Reflections were recorded after each FGD to capture and document valuable insights shared during the discussions. The duration of the FGD ranged from 50 minutes to 75 minutes.

Data Analysis

Interviews were audio-recorded following verbal consent and transcribed verbatim. Transcripts were checked for validity, and any identified errors were corrected. Two coders (SY and HG) reviewed the transcripts independently and thematically analyzed the data using NVivo (version 12; Lumivero). We used reflexive thematic analysis following each completed interview [19], contributing to the ongoing refinement and direction of the interview guide for subsequent interviews. The coding categories evolved from initial open coding to more analytical coding of the text, ultimately revealing a series of interconnected themes and patterns. The analysis and interviews continued until no new emerging themes were identified. In case of discrepancies, iterative discussions involving study team members were conducted to resolve any differences and ensure consistency in the analysis process.

Recognizing the inherent influence of the coders' subjective perspectives in the research process, the study team prioritized strategies aimed at effectively managing these preconceptions and upholding the integrity and credibility of the analysis. Specifically, we implemented the following measures: (1) the establishment of an elaborate coding protocol, meticulously designed to promote consistency and minimize potential subjective interpretations and (2) regular engagement in peer-debriefing sessions and member checking to validate the interpretations and enhance the credibility and confirmability of our findings.

Ethical Considerations

Ethics approval was obtained from SingHealth Centralized institutional review board (2022/2329). Prior to the FGD, a proficient research coordinator (HG) engaged with each participant individually to meticulously review the participant information sheet and consent form. Particular emphasis was placed on elucidating the study's objectives, along with an extensive exploration of potential foreseeable risks and benefits. Upon satisfactory comprehension, the participants were invited to provide their informed consent by endorsing the documentation. Furthermore, they were duly informed that all collected data would undergo a stringent deidentification process to preserve anonymity. To uphold transparency and equity in the compensation process, the participants were explicitly notified beforehand that no form of compensation would be provided for their involvement in this study. It was reiterated that participation in the study was entirely voluntary.

Results

Characteristics of Participants

In total, 18 clinicians were contacted, and 13 responded positively to the invitation and participated in the FGDs. The remaining 5 clinicians declined to participate, citing time constraints and lack of interest as the reasons. The FGDs were conducted based on participants' seniority at work. Most participants were male (n=8, 61%) and aged between 31 and 50 years (n=12, 92%). The participants held various designations in their respective roles, including resident (n=3, 23%), associate consultant (n=1, 8%), consultant (n=5, 38%), and senior consultant (n=4, 31%). Notably, more than three-quarters (n=10, 77%) of the participants had no prior experience with AI-enabled CDSSs. Detailed characteristics of the study participants are shown in Table 1.

XSL•FO
RenderX

Table 1. Characteristics of participants (N=13).

Characteristics	Participants, n (%)
Age (years)	
31-40	6 (46)
41-50	6 (46)
Older than 50 years	1 (8)
Sex	
Female	5 (38)
Male	8 (61)
Seniority at work	
Resident	3 (23)
Associate consultant	1 (8)
Consultant	5 (38)
Senior consultant	4 (31)
Prior experience with AI[a]-enabled CDSSs[b]	
Yes	3 (23)
No	10 (77)

[a]AI: artificial intelligence.

[b]CDSS: clinical decision support system.

Our analysis yielded 2 themes and 9 subthemes that represented the participants' perspectives concerning the use and impact of the APA tool on patient care and clinical practice and the challenges to adoption and application of the APA tool. Descriptions of themes and subthemes are presented in Textbox 1.

Textbox 1. Main themes and subthemes.

Use and impact of the AI–enabled Prescription Advisory (APA) tool on patient care and clinical practice

- Supporting decision-making for patients with comorbidities (artificial intelligence [AI]–powered drug class recommendations)
- Facilitating doctor-patient communication (diabetic complications risk predictions)
- Enhancing clinical confidence through cross-checking (color-coded AI-generated recommendations)
- Serving as a gatekeeper against medical negligence

Challenges concerning adoption and application of the APA tool

- Interpretability issues due to the lack of standardized guidelines on AI risk predictions
- Mistrust in the system driven by perceived lack of transparency around system development and information sourcing
- Limited applicability in a specialist setting given extensive expertise and patient care accountability of endocrinologists
- Concerns about potential harm in light of occasional contradictions between the APA tool recommendation and a clinician's professional judgment
- Frustration with technical issues associated with the tool implementation

Use and Impact of APA Tool on Patient Care and Clinical Practice

When asked about their experience with the APA tool, most participants expressed a positive impact, highlighting its potential to guide clinical decision-making as a key benefit. Specifically, the participants appreciated the AI algorithm's ability to provide drug class recommendations based on patients' laboratory markers. Notably, the tool not only simplified chronic disease management but also assisted in identifying instances of suboptimal disease control that might have otherwise gone unnoticed during consultations. This streamlined approach proved invaluable in guiding clinicians toward effectively managing comorbidities and reducing the risk of long-term complications in patients with diabetes.

> *I think the most valuable part for me is the lipid control feature. Sometimes when engrossed in discussing patients' diabetes treatment plans, which is anyway their primary reasons for seeking consultation, I may overlook the assessment of their LDL-c levels. With the APA tool, a quick glance provides a clear indication of whether they are on*

XSL•FO
RenderX

> *target or not. There isn't much extra clinical information that is required by the tool, so I am able to rely on the medication recommendations to appropriately adjust the medication for hyperlipidemia.* [FGD 2, senior consultant]

Furthermore, some participants saw the AI-generated complication risk predictions as a helpful resource in "convincing patients to adhere to certain treatments or treatment plans." By presenting visible evidence regarding the potential risks linked to noncompliance or inadequate treatment, the tool showed considerable potential in facilitating doctor-patient communication based on risk prediction.

> *The complications risk prediction feature stands out as particularly beneficial to me. For example, it provides an alert regarding the risk of hypoglycemia. When the risk level is classified as moderate or high, this information helped me better persuade patients to consider specific treatments or to improve their compliance with the recommended approach.* [FGD 2, senior consultant]

Some participants pointed out the lack of quantitative representation for the AI-generated complication risk prediction scores. They proposed an interactive time series graph that would visually illustrate the fluctuations in risk scores over time following the adoption of the tool's recommendations. Participants believed that integrating visual aids would enhance patients' understanding of their current risk levels associated with complications and promote the benefits of adhering to treatment plans.

> *The ability to visually present individual risk in a quantitative way through graphical or pictorial means and illustrate the potential changes that may occur after adopting the systems' recommendations would improve information delivery. I personally believe that patients are more inclined to accept the recommendations when they see their risk in a pictorial or a graphical format.* [FGD 1, associate consultant]

The color-coded AI-generated recommendations served as an additional point of reference during consultations, particularly when discrepancies emerged between the tool's recommendations and the clinician's own knowledge. This feature not only fostered critical thinking but also prompted clinicians to consider additional clinical histories that might have been overlooked initially. Overall, clinicians reported an enhanced level of confidence in their clinical decisions, thereby "improving the quality of patient care."

> *So, the tool helps to reinforce my decision-making. The color-coded recommendations provide a clear visual indication, prompting me to address any discrepancies that may arise between the tool's suggestions and my own clinical plan. In this case, I delve into additional clinical histories that the tool does not have access to and elucidate the rationale behind my decisions. This process enhances my confidence and guides better decision-making during the clinical visit, which can improve the quality of patient care.* [FGD 1, consultant]

By and large, the participants perceived the APA tool as a mechanism to prevent the risk of negligence, especially in fast-paced clinical environments. Acting as a "gatekeeper for patient safety," the APA tool effectively identified and flagged abnormal results, mitigating the risk of overlooking important tasks. The tool was regarded as a valuable partner in pursuit of delivering high-quality and safe care to patients.

> *I like the idea of the tool as a gatekeeper for patient safety. Making sure doctors don't forget things, reminding us to check and act on abnormal results. I think that is useful for busy clinics.* [FGD 1, associate consultant]

Challenges to the Adoption and Application of APA Tool

Although participants generally acknowledged the beneficial effects of the APA tool on quality patient care and clinical practice, they equally expressed reservations about incorporating and using the tool in their own clinical settings. One major concern centered around the interpretability of the automatically generated AI score when new drugs were recommended. While participants appreciated the availability of the scoring system to inform the likelihood of achieving treatment targets based on the recommendations, they remained unsure about the interpretability of the AI score.

> *I think it is quite interesting that the system is able to provide different percentages of achieving optimum blood pressure when different combinations of new drugs are used. However, my question is if plan A gives a score of 48 while plan B gives a score of 45, are these recommendations still clinically relevant? I mean, of course, the situation is more direct in cases with scores such as 98 and 88, then it will make more sense to pick the plan with 98% of likelihood.* [FGD 1, consultant]

A sense of mistrust in the APA system emerged, which appeared to stem from the unfamiliarity surrounding AI-based recommendations and concerns regarding transparency of the information sourcing and system development. Some participants openly expressed their hesitancy in adopting the APA tool due to the absence of essential clinical data. Without access to this information, they were not confident enough to use the tool.

> *When it [APA] was launched, a lot of us were not very sure how it was developed. I think part of the reason why we did not use it very much is also because we are not so familiar with how this system came about, what kind of information was used, and where the information came from. Is it also possible that critical information was not captured in the system? I can't trust totally, and [I am] not confident with what I'm seeing at the moment.* [FGD 3, senior consultant]

However, participants expressed openness to embrace the tool if they were presented with additional information. They

XSL•FO
RenderX

emphasized the importance of transparent communication regarding the evidence supporting the system and the sources of information used. By gaining a clearer understanding of the logic and rules behind the recommendations, they would be more inclined to use the tool in their clinical practice.

> *[T]hat being said, if more information or transparent communication is given to us, I might be more inclined to use it in clinics. As I know the logic and rules behind these recommendations and where they are sourced from.* [FGD 3, senior consultant]

While a minority, some clinicians exhibited strong confidence in their own clinical judgment and thus did not see the necessity to rely on the APA tool. They felt that their experience and specialist training surpassed the assistance provided by the AI-driven system. In addition, they highlighted the potential ramifications of relying on CDSS recommendations, emphasizing that the responsibility for patient outcomes ultimately rested with the clinician. Consequently, this attitude led to a reluctance to use the tool, particularly among those who believed that they possessed the requisite expertise to make well-informed decisions in patient care.

> *I would say that I'm as good or even better than the system. I don't feel the need to rely on it; I'll just do what I do. We are all trained endocrinologists, so we trust our judgment because that has been our bread and butter for many years. At the end of the day, we bear the responsibility for our patients, so you know, if the algorithm makes a sound decision, but something unfortunate ever happens to the patient, then it's still our own accountability on the line.* [FGD 1, consultant]

These clinicians suggested the potential for the APA tool to bring benefits to the wider primary care community, particularly those who may be “less familiar with endocrinology clinical practices.” They believed that the tool could assist general practitioners in effectively managing patients with complex cases and improve patient engagement.

> *These recommendations would be more valuable in a primary healthcare setting, where doctors may not have extensive knowledge of clinical practices related to novel glucose-lowering medications and insulin titration, especially in complex cases. I think implementing the AP tool in such settings would greatly help doctors in improving patient engagement and care.* [FGD 1, consultant]

Another important theme was related to the potential harm of the APA tool’s drug recommendations on patients. Participants noted that the recommendation occasionally contradicted their own professional judgment. They cautioned against solely relying on algorithmic recommendations for clinical decision-making.

> *Some of the recommendations go against your clinical judgement. For example, I have two patients and the AI recommendation was to add a beta blocker to someone who doesn't have ischemic heart disease as a second line agent. That's just not something that we would normally do. So have to exercise caution too!* [FGD 1, consultant]

Finally, the participants expressed their frustration with the technical issues associated with the integration of the APA tool into the EMR system. The slow loading of clinical notes resulted in delayed clinical consultations, which added unnecessary mental burdens for some participants. Moreover, there were instances in which the clinical notes failed to load entirely, thereby affecting the quality of patient care.

> *One significant issue we encounter after implementing the CDSS is the considerable lag in loading clinical notes. It takes a few minutes to retrieve the clinical notes. So, by that time, I'm typically already engaged in a conversation with the patient, and we may even come up with a plan without the notes being available. In some instances, I can't even see the clinical notes at all.* [FDG 4, registrar]

Discussion

Principal Findings

This qualitative study explored clinicians’ perspectives on the use and impact of the APA tool, as well as challenges to its adoption and application in clinical practice. In terms of use, the APA tool offers several useful features to assist clinicians in effectively managing diabetes. As shown in the literature, patients with T2DM frequently experience multiple comorbidities, which may add complexity to pharmacotherapy management and increase the mental burden of prescribing practices [20]. Our findings suggest that the AI algorithms for drug alteration embedded in the APA tool were generally viewed as a good knowledge resource in supporting the clinician’s decision-making on drug modifications at the point of care, particularly for patients with T2DM with comorbidities.

Complications arising from diabetes pose a significant burden on the public health care system [21,22]. In light of this, an important feature developed in the APA tool was the diabetic complications risk prediction that provides information on the likelihood of developing the 6 most common diabetic complications in patients with T2DM [23,24]. We found that participants viewed the risk prediction as having a positive impact on patient care by facilitating early doctor-patient communication and initiating prompt clinical responses to delay the progression of complications associated with diabetes. This finding is similar to that of other research that AI-enabled CDSSs had a positive impact on patient-provider encounters and shared decision-making [25,26]. Therefore, appropriate use of risk prediction could enable clinicians to take early proactive measures to reduce the risk of developing diabetic complications, ultimately reducing the health care costs associated with diabetes [27,28].

Despite the perceived merits of AI-generated risk scores, the absence of clear frameworks (or the scientific basis from which recommendations were derived) limited the interpretability and usability of the risk scores and subsequent follow-up actions. This has been similarly identified in the literature as a key hindrance to clinical adoption [29,30]. As knowledge is

XSL•FO
RenderX

deciphered differently based on personal experience and beliefs, the interpretation of scores could be dependent on the subjective attitudes of clinicians in decision-making [31]. Indeed, recent research indicates that the varying levels of knowledge and self-reported behavior among clinicians affect their approach in clinical practice, leading to potential noncompliance with the system recommendations [13]. Furthermore, as shown in our study, some clinicians chose to abstain from using the APA tool entirely because of their lack of trust in the quality of model inputs and parameters, as well as their concerns regarding the logic behind the AI outputs, often referred to as the "black box" situation [32,33]. To ensure a successful expansion of the APA tool within the clinical ecosystem, more effort should be directed to obtain a better comprehension of clinicians regarding the AI technology's capabilities and the use of explainable frameworks to enhance transparency and clinician engagement [34-36].

As with the literature, clinicians in our study cautioned against being overly reliant on the APA tool, as occasional erroneous recommendations generated by the systems might prompt users to override a correct decision they have already made [26,37]. When users are subjected to automation bias, a tendency to overaccept system recommendations as a heuristic replacement of vigilant information processing [38] and medical errors ensue from following incorrect recommendations. Not only does it predispose patients to even greater harm, but it also diminishes the intention of using AI-enabled CDSSs [39]. Our results underscore the importance of collaborative intelligence, where users and AI work synergistically to enhance patient care. The human-in-the-loop concept suggests that while human oversight is active, overdependence on AI-enabled CDSSs is equally harmful. The optimal approach involves granting clinicians full control over the decision-making process while using AI to offer recommendations and inputs [40]. Clinical decisions, therefore, cannot be made without active involvement from clinicians to serve as gatekeepers, prevent negligence, and ensure patient safety. The seemingly conflicting recommendations identified in this study should be viewed as a catalyst that prompts critical thinking, and more effort should be made to confront meaningful disagreements. Also, encouraging clinicians to check on discrepancies may enhance their confidence in decision-making [41].

Finally, a significant obstacle that hindered the adoption of APA tool pertains to concerns surrounding accountability and liability, which is in line with the literature [36,42]. While ethical considerations regarding the use of AI persist, establishing well-defined clinical standards and codes of conduct for adopting APA tools can foster a culture of shared responsibility, moving away from a single form of attribution of responsibility. In addition, instead of focusing on assigning blame, it would be more constructive to acknowledge and commend clinicians' efforts in integrating the outputs of AI-enabled CDSSs into their decision-making process, as long as the adoption of recommendations adheres to clinical standards, legal obligations, and ethical principles. This approach would motivate clinicians to embrace the most advanced medical technologies available in clinical practice, even if it means having to make a judgment call.

Collectively, the findings underscore the promising impact of adopting the APA tool within clinical settings and its potential to usher in notable enhancements in health policy. While the rapid integration of AI-based CDSSs in health care has presented promising potential for improved patient outcomes and streamlined clinical workflows, the persistent liability concerns among clinicians have created a barrier to the widespread adoption of these advanced technologies. With clinicians ultimately bearing the responsibility for any medical negligence, even after consulting with AI-enabled CDSS recommendations, there arises an urgent need for a comprehensive medicolegal framework. Such a framework must emphasize the allocation of liability among users, while also ensuring transparency in the decision-making processes of these AI tools [43]. For instance, the policy should delineate clear protocols for the documentation of AI-based recommendations, ensuring that the decision-making process is well-documented and easily accessible for medicolegal reviews. In addition, it is crucial to establish standardized protocols for the continuous evaluation and improvement of AI algorithms to minimize the risk of errors and improve the accuracy of recommendations. Creating an environment that fosters trust in AI technologies through a robust medicolegal framework will ultimately encourage clinicians to embrace these tools, leading to enhanced health care delivery and patient outcomes.

Limitations

This study provides valuable insights into the benefits and adoption challenges of an AI-based CDSS in its early stage of implementation. This study has some limitations. The study participants were limited to endocrinologists in a tertiary hospital; therefore, the generalizability of the findings to other health care settings may be limited. The sample size of the study is small, which may hinder the generation of comprehensive insights that better represent the broader context. As adoption of AI technology in clinical settings is still in its early stage, assembling a large cohort of clinicians for an in-depth analysis of AI-enabled CDSS implementation can be challenging due to the limited number of early adopters. Nevertheless, our findings shed light on the initial experiences and perceptions of a key group of clinicians, offering a foundation for future research and more extensive investigations. Our study's sample size also aligns with the systematic review, which found that empirical studies, especially those with homogenous populations and narrowly defined objectives, typically achieve data saturation with 9-17 interviews [44]. Further research is needed to explore the use and impact of the APA tool in different clinical settings, such as primary care. As suggested by our participants, the use of the APA tool can be particularly beneficial to general practitioners who are responsible for managing a wide range of conditions and require access to a breadth of knowledge base across various specialty areas. Despite early findings on the APA tool's use and adoption challenges, its long-term impacts on clinical and economic outcomes remain unknown. A subsequent larger evaluation is warranted to compare the APA tool with a standard of care. Finally, we did not explore the perspectives of patients with diabetes as the important end users of the APA tool.

XSL•FO
RenderX

Incorporating their perspectives may have contributed to a richer understanding.

Conclusions

AI-enabled CDSSs, such as the APA tool, has the potential to enhance clinical practice and patient care. Clinicians found certain features such as AI algorithms on medication adjustment and complication risk predictions useful in managing patients with T2DM with comorbidities and facilitating doctor-patient communication. However, interpretability of the risk scores, concerns about overreliance and automation bias, and issues surrounding accountability and liability were commonly cited as challenges inhibiting the adoption and application of the APA tool in endocrinology clinical settings. Further work is required to address these concerns effectively to enhance the tool's acceptance and applicability in relevant contexts.

Acknowledgments

We thank all participants for their participation in this study.

Data Availability

The data that support the findings of this study are available upon reasonable request from the corresponding author. The data are not publicly available due to information that could compromise the privacy of research participants.

Authors' Contributions

SY and YMB were responsible for the conception, design, and the whole protocol of the study. SY oversaw the study. HG, PCL, HCT, MMT, DSTL, AK, CS, DC, DSS, SYTT, AJWW, CHMC, and ZW were responsible for the acquisition of study data. SY and HG were responsible for data analysis and interpretation of study data. SY and HG drafted the manuscript. All authors critiqued the output and read and approved the final manuscript.

Conflicts of Interest

None declared.

References

1. Er B, James L, Ma S, Ministry of Health. The Burden of Disease in Singapore, 1990-2017. Singapore: Ministry of Health; 2019.
2. Forbes JM, Cooper ME. Mechanisms of diabetic complications. Physiol Rev 2013;93(1):137-188 [FREE Full text] [doi: 10.1152/physrev.00045.2011] [Medline: 23303908]
3. Papatheodorou K, Banach M, Bekiari E, Rizzo M, Edmonds M. Complications of diabetes 2017. J Diabetes Res 2018;2018:3086167 [FREE Full text] [doi: 10.1155/2018/3086167] [Medline: 29713648]
4. Crowley K. Patient Education: The ABCs of Diabetes (The Basics). UpToDate. URL: https://www.uptodate.com/contents/the-abcs-of-diabetes-the-basics?search=abc%20diabetes&source=search_result&selectedTitle=1~150&usage_type=default&display_rank=1 [accessed 2024-05-23]
5. American Diabetes Association. Standards of medical care in diabetes—2016 abridged for primary care providers. Clin Diabetes 2016;34(1):3-21 [FREE Full text] [doi: 10.2337/diaclin.34.1.3] [Medline: 26807004]
6. Sim LLW, Ban KHK, Tan TW, Sethi SK, Loh TP. Development of a clinical decision support system for diabetes care: a pilot study. PLoS One 2017;12(2):e0173021 [FREE Full text] [doi: 10.1371/journal.pone.0173021] [Medline: 28235017]
7. Chen R, Jiang HQ, Huang C, Bau C. Clinical decision support system for diabetes based on ontology reasoning and TOPSIS analysis. J Healthc Eng 2017;2017:4307508 [FREE Full text] [doi: 10.1155/2017/4307508] [Medline: 29312655]
8. Jia P, Zhao P, Chen J, Zhang M. Evaluation of clinical decision support systems for diabetes care: an overview of current evidence. J Eval Clin Pract 2019 Feb;25(1):66-77. [doi: 10.1111/jep.12968] [Medline: 29947136]
9. Chien T, Ting H, Chen C, Yang C, Chen C. A clinical decision support system for diabetes patients with deep learning: experience of a Taiwan medical center. Int J Med Sci 2022;19(6):1049-1055 [FREE Full text] [doi: 10.7150/ijms.71341] [Medline: 35813300]
10. Sutton RT, Pincock D, Baumgart DC, Sadowski DC, Fedorak RN, Kroeker KI. An overview of clinical decision support systems: benefits, risks, and strategies for success. NPJ Digit Med 2020;3:17 [FREE Full text] [doi: 10.1038/s41746-020-0221-y] [Medline: 32047862]
11. Contreras I, Vehi J. Artificial intelligence for diabetes management and decision support: literature review. J Med Internet Res 2018;20(5):e10775 [FREE Full text] [doi: 10.2196/10775] [Medline: 29848472]
12. Spänig S, Emberger-Klein A, Sowa J, Canbay A, Menrad K, Heider D. The virtual doctor: an interactive clinical-decision-support system based on deep learning for non-invasive prediction of diabetes. Artif Intell Med 2019;100:101706. [doi: 10.1016/j.artmed.2019.101706] [Medline: 31607340]

XSL•FO
RenderX

13. Nimri R, Tirosh A, Muller I, Shtrit Y, Kraljevic I, Alonso MM, et al. Comparison of insulin dose adjustments made by artificial intelligence-based decision support systems and by physicians in people with type 1 diabetes using multiple daily injections therapy. Diabetes Technol Ther 2022;24(8):564-572. [doi: 10.1089/dia.2021.0566] [Medline: 35325567]
14. Bajgain B, Lorenzetti D, Lee J, Sauro K. Determinants of implementing artificial intelligence-based clinical decision support tools in healthcare: a scoping review protocol. BMJ Open 2023;13(2):e068373. [doi: 10.1136/bmjopen-2022-068373] [Medline: 36822813]
15. Tong A, Sainsbury P, Craig J. Consolidated criteria for reporting qualitative research (COREQ): a 32-item checklist for interviews and focus groups. Int J Qual Health Care 2007;19(6):349-357. [doi: 10.1093/intqhc/mzm042] [Medline: 17872937]
16. Sun X, Bee YM, Lam SW, Liu Z, Zhao W, Chia SY, et al. Effective treatment recommendations for type 2 diabetes management using reinforcement learning: treatment recommendation model development and validation. J Med Internet Res 2021;23(7):e27858 [FREE Full text] [doi: 10.2196/27858] [Medline: 34292166]
17. Wan Q, Makeham M, Zwar NA, Petche S. Qualitative evaluation of a diabetes electronic decision support tool: views of users. BMC Med Inform Decis Mak 2012;12:61 [FREE Full text] [doi: 10.1186/1472-6947-12-61] [Medline: 22759239]
18. Pratt R, Saman DM, Allen C, Crabtree B, Ohnsorg K, Sperl-Hillen JM, et al. Assessing the implementation of a clinical decision support tool in primary care for diabetes prevention: a qualitative interview study using the Consolidated Framework for Implementation Science. BMC Med Inform Decis Mak 2022;22(1):15 [FREE Full text] [doi: 10.1186/s12911-021-01745-x] [Medline: 35033029]
19. Braun V, Clarke V. Using thematic analysis in psychology. Qual Res Psychol 2006;3(2):77-101. [doi: 10.1191/1478088706qp063oa]
20. Foo KM, Sundram M, Legido-Quigley H. Facilitators and barriers of managing patients with multiple chronic conditions in the community: a qualitative study. BMC Public Health 2020;20(1):273 [FREE Full text] [doi: 10.1186/s12889-020-8375-8] [Medline: 32106838]
21. McEwen LN, Herman WH. Health care utilization and costs of diabetes. In: Cowie CC, Casagrande SS, Menke A, Cissell MA, Eberhardt MS, Meigs JB, editors. Diabetes in America. Bethesda, MD: National Institute of Diabetes and Digestive and Kidney Diseases; 2018.
22. Disdier-Flores OM, Rodríguez-Lugo LA, Pérez-Perdomo R, Pérez-Cardona CM. The public health burden of diabetes: a comprehensive review. P R Health Sci J 2001;20(2):123-130. [Medline: 11561471]
23. Deshpande AD, Harris-Hayes M, Schootman M. Epidemiology of diabetes and diabetes-related complications. Phys Ther 2008;88(11):1254-1264 [FREE Full text] [doi: 10.2522/ptj.20080020] [Medline: 18801858]
24. Sagoo MK, Gnudi L. Diabetic nephropathy: an overview. Methods Mol Biol 2020;2067:3-7. [doi: 10.1007/978-1-4939-9841-8_1] [Medline: 31701441]
25. Labinsky H, Ukalovic D, Hartmann F, Runft V, Wichmann A, Jakubcik J, et al. An AI-powered clinical decision support system to predict flares in rheumatoid arthritis: a pilot study. Diagnostics (Basel) 2023;13(1):148 [FREE Full text] [doi: 10.3390/diagnostics13010148] [Medline: 36611439]
26. Romero-Brufau S, Wyatt K, Boyum P, Mickelson M, Moore M, Cognetta-Rieke C. A lesson in implementation: a pre-post study of providers' experience with artificial intelligence-based clinical decision support. Int J Med Inform 2020;137:104072. [doi: 10.1016/j.ijmedinf.2019.104072] [Medline: 32200295]
27. Lloyd A, Sawyer W, Hopkinson P. Impact of long-term complications on quality of life in patients with type 2 diabetes not using insulin. Value Health 2001;4(5):392-400 [FREE Full text] [doi: 10.1046/j.1524-4733.2001.45029.x] [Medline: 11705130]
28. Trikkalinou A, Papazafiropoulou AK, Melidonis A. Type 2 diabetes and quality of life. World J Diabetes 2017;8(4):120-129 [FREE Full text] [doi: 10.4239/wjd.v8.i4.120] [Medline: 28465788]
29. Ching T, Himmelstein DS, Beaulieu-Jones BK, Kalinin AA, Do BT, Way GP, et al. Opportunities and obstacles for deep learning in biology and medicine. J R Soc Interface 2018;15(141):20170387 [FREE Full text] [doi: 10.1098/rsif.2017.0387] [Medline: 29618526]
30. Rudin C. Stop explaining black box machine learning models for high stakes decisions and use interpretable models instead. Nat Mach Intell 2019;1(5):206-215 [FREE Full text] [doi: 10.1038/s42256-019-0048-x] [Medline: 35603010]
31. Laws RA, Kemp LA, Harris MF, Davies GP, Williams AM, Eames-Brown R. An exploration of how clinician attitudes and beliefs influence the implementation of lifestyle risk factor management in primary healthcare: a grounded theory study. Implement Sci 2009;4:66 [FREE Full text] [doi: 10.1186/1748-5908-4-66] [Medline: 19825189]
32. Amann J, Blasimme A, Vayena E, Frey D, Madai VI, Precise4Q consortium. Explainability for artificial intelligence in healthcare: a multidisciplinary perspective. BMC Med Inform Decis Mak 2020;20(1):310 [FREE Full text] [doi: 10.1186/s12911-020-01332-6] [Medline: 33256715]
33. Shortliffe EH, Sepúlveda MJ. Clinical decision support in the era of artificial intelligence. JAMA 2018;320(21):2199-2200. [doi: 10.1001/jama.2018.17163] [Medline: 30398550]
34. Cutillo CM, Sharma KR, Foschini L, Kundu S, Mackintosh M, Mandl KD, MI in Healthcare Workshop Working Group. Machine intelligence in healthcare-perspectives on trustworthiness, explainability, usability, and transparency. NPJ Digit Med 2020;3:47 [FREE Full text] [doi: 10.1038/s41746-020-0254-2] [Medline: 32258429]

XSL•FO
RenderX

35. Reddy S. Explainability and artificial intelligence in medicine. Lancet Digit Health 2022;4(4):e214-e215 [FREE Full text] [doi: 10.1016/S2589-7500(22)00029-2] [Medline: 35337639]
36. Jones C, Thornton J, Wyatt J. Artificial intelligence and clinical decision support: clinicians' perspectives on trust, trustworthiness, and liability. Med Law Rev 2023;31(4):501-520 [FREE Full text] [doi: 10.1093/medlaw/fwad013] [Medline: 37218368]
37. Goddard K, Roudsari A, Wyatt JC. Automation bias: a systematic review of frequency, effect mediators, and mitigators. J Am Med Inform Assoc 2012;19(1):121-127 [FREE Full text] [doi: 10.1136/amiajnl-2011-000089] [Medline: 21685142]
38. Skitka L, Mosier K, Burdick M. Does automation bias decision-making? Int J Hum Comput Stud 1999;51(5):991-1006 [FREE Full text] [doi: 10.1006/ijhc.1999.0252]
39. Mosier K, Skitka L. Human Decision Makers and Automated Decision Aids: Made for Each Other?. Mahwah, NJ: Erlbaum; 1996:201-220.
40. Haselager P, Schraffenberger H, Thill S, Fischer S, Lanillos P, van de Groes S, et al. Reflection machines: supporting effective human oversight over medical decision support systems. Camb Q Healthc Ethics 2023:1-10. [doi: 10.1017/S0963180122000718] [Medline: 36624620]
41. Muhiyaddin R, Abd-Alrazaq AA, Househ M, Alam T, Shah Z. The impact of clinical decision support systems (CDSS) on physicians: a scoping review. Stud Health Technol Inform 2020;272:470-473. [doi: 10.3233/SHTI200597] [Medline: 32604704]
42. Funer F, Liedtke W, Tinnemeyer S, Klausen AD, Schneider D, Zacharias HU, et al. Responsibility and decision-making authority in using clinical decision support systems: an empirical-ethical exploration of German prospective professionals' preferences and concerns. J Med Ethics 2023;50(1):6-11 [FREE Full text] [doi: 10.1136/jme-2022-108814] [Medline: 37217277]
43. Garell C, Svedberg P, Nygren JM. A legal framework to support development and assessment of digital health services. JMIR Med Inform 2016;4(2):e17 [FREE Full text] [doi: 10.2196/medinform.5401] [Medline: 27226391]
44. Hennink M, Kaiser B. Sample sizes for saturation in qualitative research: a systematic review of empirical tests. Soc Sci Med 2022;292:114523 [FREE Full text] [doi: 10.1016/j.socscimed.2021.114523] [Medline: 34785096]

Abbreviations

AI: artificial intelligence
APA: AI-enabled Prescription Advisory
CDSS: clinical decision support system
COREQ: Consolidated Criteria for Reporting Qualitative Research
EMR: electronic medical record
FGD: focus group discussion
T2DM: type 2 diabetes mellitus

Edited by A Kushniruk; submitted 17.07.23; peer-reviewed by D Heider, N Minian; comments to author 03.11.23; revised version received 07.11.23; accepted 05.05.24; published 13.06.24.

Please cite as:
Yoon S, Goh H, Lee PC, Tan HC, Teh MM, Lim DST, Kwee A, Suresh C, Carmody D, Swee DS, Tan SYT, Wong AJW, Choo CHM, Wee Z, Bee YM
Assessing the Utility, Impact, and Adoption Challenges of an Artificial Intelligence–Enabled Prescription Advisory Tool for Type 2 Diabetes Management: Qualitative Study
JMIR Hum Factors 2024;11:e50939
URL: https://humanfactors.jmir.org/2024/1/e50939
doi:10.2196/50939
PMID:38869934

©Sungwon Yoon, Hendra Goh, Phong Ching Lee, Hong Chang Tan, Ming Ming Teh, Dawn Shao Ting Lim, Ann Kwee, Chandran Suresh, David Carmody, Du Soon Swee, Sarah Ying Tse Tan, Andy Jun-Wei Wong, Charlotte Hui-Min Choo, Zongwen Wee, Yong Mong Bee. Originally published in JMIR Human Factors (https://humanfactors.jmir.org), 13.06.2024. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on https://humanfactors.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Preferences, Needs, and Values of Patients With Chronic Obstructive Pulmonary Disease Attending a Telehealth Service: Qualitative Interview Study

Camilla Wong Schmidt[1,2], MSc; Karen Borgnakke[2], MA, PhD, DHA; Anne Frølich[3,4], MD, PhD; Lars Kayser[2], MD, PhD

[1]Medical Department, Holbæk Sygehus, Region Zealand, Holbæk, Denmark

[2]Section of Health Services Research, University of Copenhagen, Copenhagen, Denmark

[3]Innovation and Research Centre for Multimorbidity, Slagelse Hospital, Region Zealand, Slagelse, Denmark

[4]Department of Public Health, University of Copenhagen, Copenhagen, Denmark

Corresponding Author:
Camilla Wong Schmidt, MSc
Medical Department
Holbæk Sygehus
Region Zealand
Smedelundsgade 60
Holbæk, 4300
Denmark
Phone: 45 59484000
Email: cws@sund.ku.dk

Abstract

Background: Digitally assisted health care services and technologies are gaining popularity. They assist patients in managing their conditions, thereby reducing the burden on health care staff. Digital health care enables individuals to receive care that is more tailored to their needs and preferences. When implemented properly, it can promote equity by considering each person's opportunities and limitations in the context of health care needs, preferences, values, and capabilities.

Objective: This study aims to understand the needs, values, and preferences of individuals with chronic obstructive pulmonary disease (COPD) who are provided with a 24/7 digital health care service. Furthermore, we aim to understand the dynamics of the communities to which they belong and how these communities intersect. This will provide us with the essential knowledge to establish new methods of providing education, including the development of educational activities for health professionals to engage, train, and empower people living with COPD.

Methods: The study included 7 informants diagnosed with COPD who received 24/7 digital health care service support from a regional project in Region Zealand, Denmark. The informants were visited 4 times during 2 months, including a "Hello" visit, a day with a semistructured interview, and 2 days with field observations. The informants participated in a semistructured interview, following participant observation and an ethnographic approach. The interview content was analyzed using an inductive methodology to categorize the empirical data.

Results: Using the inductive approach, we identified 3 main categories related to the informants' needs, values, and preferences: (1) Health, (2) Value Creation, and (3) Resources. These 3 main categories were based on 9 subcategories: (1) health and barriers, (2) self-monitoring, (3) medication, (4) behavior, (5) motivation, (6) hobbies, (7) social networks, (8) health professionals, and (9) technology. These findings revealed that the informants placed value on maintaining their daily activities and preserving their sense of identity before the onset of COPD. Furthermore, they expressed a desire not to be defined by their COPD, as conversations about COPD often shifted away from the topic.

Conclusions: Digital health solutions and the health care professionals who offer them should prioritize the individuals they serve, considering their needs, values, and preferences rather than solely focusing on the medical condition. This approach ensures the highest level of daily living and empowerment for those living with long-term health conditions. The communities surrounding individuals must engage in constant interaction and collaboration. They should work together to incorporate people's needs, values, and preferences into future digital health services, thereby promoting empowerment and self-management. New educational

XSL•FO
RenderX

programs aimed at developing the digital health service competencies of registered nurses should facilitate collaboration between the 2 communities. This collaboration is essential for supporting patients with long-term health conditions in their daily activities.

(JMIR Hum Factors 2024;11:e53131) doi:10.2196/53131

KEYWORDS

people with long-term health condition; patient education; COPD; digital health; ethnography; inductive; ethnographic; chronic; lung; lungs; pulmonary; respiratory; self-management; interview; interviews; qualitative; experience; experiences; attitude; attitudes; opinion; perception; perceptions; perspective; perspectives; acceptance

Introduction

Background

In recent years, a transformation has been occurring with the increased use of digitally assisted health care services and technologies. These advancements aim to reduce the burden on the health care work force by enabling patients to better manage their own conditions [1]. Digitally based health care offers an opportunity for personalized and tailored health care services that better meet the needs of individual patients. Digitalization can also help reduce inequities if introduced thoughtfully, with an awareness of both the opportunities and barriers for individuals, considering their health care needs, preferences, values, and capabilities [1].

When health care professionals have the appropriate knowledge about these factors and are trained to address them, it can facilitate meaningful conversations and better connections with those they serve. This, in turn, can increase patient motivation and ease their access to digital services [2,3]. This necessitates educational programs for health professionals that focus on understanding their patient's needs and values, capability, person-centered services, self-management, and communities surrounding the patient. These programs should be based on evidence from empirical data obtained through interviews and observations of individuals with firsthand experience using digitally enabled health care services [4].

Digital Health Care and Chronic Obstructive Pulmonary Disease

Chronic obstructive pulmonary disease (COPD) is a leading cause of death, responsible for over 3 million deaths worldwide in 2019 [5]. In Denmark, 3355 individuals died from COPD in 2020, making it the second most common cause of death in the nation [6]. As a result of the progressive loss of pulmonary function, people with COPD experience impairments in their daily activities. These impairments can inhibit mobility, leading to a sedentary lifestyle [7]. The Global Initiative for Chronic Obstructive Lung Disease (GOLD), established in 1998, has developed a set of recommendations for managing COPD. Evidence shows that self-management improves outcomes for patients with COPD and reduces the likelihood of hospitalization [8,9]. The recommendations by GOLD also address how increased self-management can help motivate and engage people, leading to positive adaptations in their health behaviors.

Digital health monitoring is one of the most widely used tools for self-management of COPD and other long-term health conditions (LTHCs) and appears to reduce the risk of both hospitalization and acute visits [10]. In 2015, a randomized controlled study was conducted with 100 people with COPD, with 48 randomly selected for home monitoring and 52 for usual care. The study found that people with COPD who used a home monitoring kit for 6 months had improved health-related quality of life and reduced anxiety scores compared with people who received usual care. The study also showed that people with home monitoring kits had fewer and shorter hospitalizations than those receiving usual care [11]. Furthermore, people with COPD who participate in telemedicine-based interventions feel safer, more empowered, and more in control of their own disease [11].

People with LTHCs who possess an enhanced ability to manage their condition themselves and experience a higher level of empowerment tend to have a reduced risk for hospitalization [12]. A likely explanation for this can be found in a qualitative study from 2017, which identified 3 main themes for individuals participating in a telemedicine intervention: (1) a sense of improved security and control; (2) a better understanding of their disease; and (3) the benefits of virtual conversations [13]. However, these studies do not take place within person-centered digital health communities and do not address the significance of the participants' sociodemographic characteristics. Evidence of sociodemographic characteristics is important, as research indicates that resourceful individuals benefit more from available health care, and those with an interest in technology may derive greater benefits from digital health solutions [14]. In addition, the prevalence of COPD and other chronic conditions, such as ischemic heart disease and type 2 diabetes mellitus, is higher in areas with populations characterized by lower sociodemographic status than in the average population [15,16]. This may contribute to the risk of inequity, as lower sociodemographic status often correlates with both lower levels of education and lower digital health literacy [15,17].

Educational Programs for the Digital Transformation

Another problem often overlooked concerning disadvantaged individuals living with 1 or more LTHCs is their reluctance to attend traditional educational services structured around scholastic planning. To reduce inequity, there is a need to develop new approaches to include this segment of the population, utilizing educational methods that are not scholastic or built on classical teaching methods such as classroom-based education. In response to this need, we have initiated a project where, based on an ethnographic approach involving interviews, observations, and the cocreation of educational materials, we will develop a new methodology inspired by social learning theory [18,19]. To provide guidance for the development of new educational programs and curricula tailored for digital transformation, we have examined, on an individual level, the

XSL•FO
RenderX

preferences, values, and needs of people with COPD in the context of unrestricted access to a person-centered, digitally enabled health service available 24/7. Informed by social learning theories, we have also explored the existence of communities these individuals are part of in relation to their everyday lives and their ability to identify potential sources of support.

The purpose of this study is to gather essential information about individuals living with COPD within the context of accessing support from a regular 24/7 digital health care service and the dynamics of the communities to which they belong. This information will empower us to design new methods for providing education, including the development of educational activities for health professionals. These activities will enable professionals to effectively engage, train, and empower people with COPD.

This has led us to the following research questions:

- Research question 1: What matters for people with COPD with respect to their needs, values, and preferences in the context of using a 24/7 digital health service?
- Research question 2: What is the role of the health care and social networks, respectively, and how are these a potential part of communities formed around the patient's health condition?

Methods

Design

This report is a part of a larger PhD project and constitutes the first of 4 articles. The overarching aim of this work is to obtain insights into the lives of people with COPD, supported by a 24/7 digital health service, and to use this information to develop a patient-case–based curriculum to educate health professionals on effectively engaging, enabling, and empowering individuals living with COPD.

The first study reported here was conducted from August 1, 2020, to January 31, 2021. This period coincided with the COVID-19 pandemic in Denmark before vaccinations were introduced to protect against severe cases. The study is qualitative in nature, inspired by ethnographic research methods, and includes semistructured interviews and field observations [20]. The field study involved visiting the homes of people with COPD 4 times over 2 months. The visits included a preliminary "Hello" visit, a day dedicated to a semistructured interview, and 2 days focused on field observations. This article focuses on the data and results gathered during the second visit, which involved conducting the semistructured interview.

Context

The digital health service utilized in this study is provided by an innovation project in Region Zealand, Denmark, called PreCare. This project is built on the Epital Care Model (ECM) [21,22]. The ECM, developed in 2016, offers a 24/7 digital health care service where individuals with LTHCs monitor their own health with the assistance of nurses from a response and coordination center (RCC). The ECM consists of 6 stages: citizens with unknown LTHCs, active and independent living, virtual assisted living, virtual assisted living with support from health care professionals, outpatient care at home, and admission to a local health clinic or hospital. It serves as a template for digital health services based on patients' medical needs [22].

In total, approximately 400 participants with COPD or ischemic heart disease were enrolled in the PreCare project over a 4-year period. At any given time, there are approximately 150 participants. Each participant was provided with a tele-home monitoring kit, which included a tablet. Additionally, for participants with COPD, the kit included a spirometer, a thermometer, a pulse oximeter, and a box containing acute medicine. The participants were supported by an RCC, which was staffed with registered nurses (RNs) and an eDoctor. According to the project protocol, participants monitored their condition daily. They could always call the RCC to discuss their condition, and the RCC regularly initiated contact to ensure participants felt safe and confident. During these conversations, self-management was also supported (C Schmidt, MSc, personal communication, 2020). In the event of deterioration, the tablet would indicate a yellow or red code and send a message to the RCC. The nurses would then respond to the code and call the participant to follow-up on the reported condition. If needed, participants can take medicine from the box to treat exacerbations. In cases of further need, the nurses would contact the eDoctor [23].

Informants

The PreCare project initially provided a list of 15 participants diagnosed with COPD, each with varying degrees of severity, all of whom expressed interest in participating in the research related to the PreCare project. Subsequently, over a 4-month period, 10 of these participants were contacted by phone, selected from the top of the list. After receiving oral information about the project, 8 of these agreed to participate. However, 2 participants were not interested and the remaining 5 were not contacted as the recruitment period had exceeded. Furthermore, 1 potential informant expressed disinterest in participating after the initial meeting. The selected number of informants was determined by the limitations of the study design. The informants were invited in 3 separate periods: 3 informants were invited from August to September for the first period, 2 informants were invited from October to November for the second period, and 2 more were invited from December to January for the third period. In total, 7 informants were recruited. After obtaining oral consent via phone, further information was sent by email to 3 informants, while the other 4 did not require this. After 1 week, all 7 informants were contacted again by phone to schedule the first in-person meeting, which took place within 1-2 weeks. This study's inclusion and exclusion criteria followed the PreCare protocol [23].

Data Collection

To establish an emotional and trustworthy relationship with the informants, we scheduled a visit to their homes (the initial visit). This approach aimed to strengthen the connection between the researcher and the informants, fostering an informal and friendly atmosphere during the interview. To conduct the semistructured interview, we used an interview guide inspired by Spradley's [20] ethnographic interview techniques. The interview was

XSL•FO
RenderX

conducted in a friendly and casual approach, allowing the informants to share their experiences and discuss their everyday lives with a chronic condition as they deemed appropriate [20,24]. All interviews were conducted in the informants' homes and were audio-recorded with their consent. As a result of the informants' background, the interviews were conducted in Danish, and only quotes were later translated into English by the first author (CWS). All informants participated in the interviews; 1 participant had his spouse present during the interview.

Interview Guide

The interview guide was developed based on sociotechnical ecosystem thinking, our concept of technology readiness, and an attempt to identify how individuals belong to 1 or more communities, inspired by social learning theory (Multimedia Appendix 1) [25-28]. We conducted the interview with an open-minded approach, including "how" questions, to enable the informants to respond as they found suitable. The interview guide was structured around 6 thematic areas: daily activities, health, measurements, communities, RCC and PreCare, and literacies and digital literacies. For each of the thematic areas, we included 1 main question and underlying questions to sustain the conversation throughout the interview. For example, the theme "daily activities" included the main question: "Can you tell me how a typical day is for you?" In the theme "health," the main question was "Can you tell me how COPD has affected your life?." The 6 themes were defined by the authors and were written in Danish.

Data Analysis

The interviews were conducted, transcribed, and analyzed by the author CWS. The transcripts were analyzed using content analysis, a method for systematically and objectively describing and quantifying phenomena. An inductive approach was used, beginning with open coding to create categories, followed by abstraction to generate main categories [29]. A 3-step content analysis was used to identify the main categories.

Analysis of Interviews

Each interview was transcribed and carefully reviewed to understand the context of the data. Subsequently, the transcripts were uploaded and coded using NVivo 12 (QSR International) [30] by CWS. Over 700 codes were identified and categorized into 66 subcategories. These subcategories were then merged to create an affinity diagram initially using paper and later repeated using NVivo. This process resulted in 9 categories, each containing 4-12 subcategories, respectively. The category "Self-Monitoring" had the fewest subcategories, while "Health Professionals and Social Network" had the most subcategories. The 9 categories were analyzed by CWS and the last author (LK) to synthesize the data into 3 main categories. CWS, who holds an MSc degree in health informatics and has been educated in qualitative methods, collaborated with LK, a professor in health service research with experience in both qualitative and quantitative analyses, for this process.

The 3 main categories identified were health, value creation, and resources (Textbox 1 and Multimedia Appendix 2). The category of health consisted of 3 subcategories: health and barriers, medication, and measurements. The category of value creation was formed from hobbies, behavior, and motivation. The category of resources was merged from 3 subcategories: social networks, health professionals, and technology. In our analysis, we paid particular attention to what matters to people with COPD, supported by the theories upon which the interview guide was built.

Textbox 1. Overview of the 3 main categories and subcategories.

1. Health

- Health and barriers
- Self-monitoring
- Medication

2. Value Creation

- Behavior
- Motivations
- Hobbies

3. Resources

- Social network
- Health professionals
- Technology

Ethical Consideration

Information regarding the study, partnerships, and data handling complies with the Helsinki Declaration and was communicated to the informants in both written and oral forms. They were informed that their participation was voluntary and anonymous and that they could revoke their consent at any time. Furthermore, they were assured that their involvement would not prevent them from participating in the PreCare project. All consent was obtained before the interview, through the signing

XSL•FO
RenderX

of a consent form. The Danish National Center for Ethics was not required to approve the study as no biological material was used. Any data obtained from the informants were treated as personal health information and handled in accordance with Danish legislation (General Data Protection Regulation [GDPR]) and securely stored on drives. Health science questionnaire surveys and interview studies that do not involve human biological material [section 14(2) of the Danish Act on Committees] do not require reporting or approval from the Danish National Centre for Ethics [31].

Results

Characteristics of the Informants

A total of 4 men and 3 women participated in the interviews (age range 52-81 years). Two informants lived with their spouses. Despite having had COPD for an average of more than 2 years, the severity of each participant's COPD varied. Some participants continued to smoke daily despite being aware of the health risks. One male participant was unable to provide information to categorize his level of education, 2 had only completed elementary school, while 4 had completed higher education. There was no evidence of their usage of technology, such as websites and participation in online communities, in relation to their medical concerns. All informants had been included in the PreCare Project for more than 6 months.

The Three Main Categories

The main categories and subcategories identified in the content analysis provide insight into and offer a comprehensive understanding of the daily life situations and experiences of the informants living with COPD. Upon reviewing these categories, attention is drawn to both the specific consequences of a COPD diagnosis and how practical hurdles and activity levels are affected in the daily lives of the informants. These impacts are described in the interviews as limitations on activities the informants were accustomed to participating in, as well as a determination to carry out specific household duties despite a decreased energy level. The duality between "restrictions" and "experimental salvage" is evident in the category of activity but is also observed when interviews approach questions such as self-monitoring. Here, they take on different meanings, tasked with reclaiming self-discipline and control on one hand, while also being concerned that daily measurements can serve as a reminder of one's limitations, akin to "being reminded of having a chronic disease." Thus, through the interviews, it becomes apparent how the informants encounter difficulties and impediments in carrying out daily tasks due to their condition. In everyday life, this translates to tasks that were once feasible but now being difficult or impossible to complete. The distinction between "then" and "now" is frequently referenced, highlighting the contrast between the condition "before I got COPD" and "the situation as a chronic."

Health

Health and Barriers

This category describes the experience of living with COPD, detailing how it has impacted daily life and outlining the physical and mental barriers experienced throughout the day or in general.

The informants did not express interest in delving deeper into their everyday lives with COPD. Instead, they prioritized discussing other aspects of their daily life or past experiences. They responded quickly to questions about COPD and then redirected the conversation toward other topics. This deliberate redirection indicated their reluctance to discuss their chronic illness.

> *Interviewer: ...Can you tell us how your diagnosis has affected your life?*
>
> *M3: So, I'm crushed. One positive thing is that I had to sell my motorcycle and all my stuffs, I used to gather a lot. We had a 400 kvm house with basement and ceiling, which was filled with enamel sign, books, magazines, tech cars and bicycles...*

The informant swiftly and effectively shifts his focus away from negative thoughts about COPD's interference and begins discussing his previous interest in used objects. He demonstrates a clear desire not to dwell on the negative aspects of his existence, opting instead to redirect the conversation toward something positive and reminiscent of happier times.

The limitations imposed by COPD forced the informants to forego certain daily activities, some of which could have contributed to an improved quality of life. The frequent shortness of breath and coughing prevented them from engaging in activities such as walking outside, performing household duties, or general personal care needs.

> *Interviewer: ...but is there other things COPD had done, that you cannot do anymore, completely?*
>
> *M1: Well, I cannot go to the city and get me a cup of coffee at the street restaurant.*
>
> *Interviewer: No, that's true...*
>
> *M1: I'm not even sure I'd be able to go to the garbage cans anymore (coughing), but when my friend comes and the weather is good (...) he drives me in that wheelchair over there, and then we sit together and drink a cup of coffee and talk...*

The informant's worry about his capacity to take out the trash underscores the profound impact that COPD has on his life, to the extent that he feels unable to leave the house without assistance. Conversations with the informant were replete with stories where social interactions played a significant role.

For some individuals, participating in a community became challenging due to shortness of breath caused by COPD. Additionally, for others, COPD had led to the complete exclusion of previous acquaintances.

> *M3 wife: In return, you have thought about how many of them you have helped (...), you don't really hear from them anymore, because now you can't help them anymore.*
>
> *M3: Yes, there are many of those whom I have been calling, "Great that we are talking to you, we were just thinking of you, by the way we have a locker that*

XSL•FO
RenderX

doesn't work". You never heard from these people again, and I have been discussing this with others, and it is true...

The informant noted that his inability to visit friends anymore, coupled with their failure to reciprocate, has made it increasingly difficult for him to maintain relationships with them. This situation has surprised him, particularly because he is no longer able to provide assistance, as reported by his wife.

Self-Monitoring

This category highlights how the informants manage and self-monitor their condition. Furthermore, it explores how the outcomes of their monitoring efforts may impact their day and their motivation to engage in activities.

In the informants' descriptions of their daily lives, the topic of self-monitoring for the PreCare project was not initially mentioned. It was only during the conversation around this subject that the activity itself was explained and, in some cases, mentioned.

M1: Yes but, it's not interesting, no(...) and then I hope in the end it can help other people too. So, I take it with pleasure, but I could still think of something more exciting things to do...

W1: Yes, but they have changes it (pause), I'm just going to write something today, I'm not in for it. It's not correct anyway (temperature)

The self-monitoring is described here as uninteresting, with 1 informant considering it a waste of time because he could find more engaging activities to do during his challenging day. The second informant emphasizes the importance of accurate measurements for individuals to actively participate in self-monitoring.

The outcomes of self-monitoring had a significant impact on the informants. The results were displayed as 1 of 3 colors—red, yellow, or green—on the tablet's display. The meaning of the color had a tangible effect on the informant's day.

Interviewer: Yes, exactly when, but then how? Because now you said that you had a red measurement yesterday was it then a difficult day when you have a red measurement?

W2: Yes, that is a stupid day, at first the mood is going down, and I am going, well yes I usually get restless, because I can't, because a day like that, I am thinking about...is it now it's going in the wrong direction...

The informant faces difficulties getting through the day when the red color appears, disrupting their daily routine. Additionally, the informant begins to question whether their COPD-related health is deteriorating or if the red color indicates a negative trend.

Conversely, the green color holds significance for the informant, particularly in contrast to the red color. Seeing a green measurement might enhance the informant's enjoyment of daily life, especially if it has been a while since they last saw a green result.

W1: It's green! (happy/excited)

W1: It haven been that for a long time....

Although there was considerable excitement surrounding the green measurement, its implications for the maintenance of the day remained unclear. However, the informant did clarify early in the conversation that she felt more motivated to venture outside into the garden on good days.

Medication

Being chronically ill entails the necessity of medication and its management which, for most of the informants, has become integrated into daily life. This category elucidates how the informants handle their medication and who supports them in managing it. The informants varied in their approaches to and understanding of medication, and the availability of help and support was crucial.

W2: Thus, those prescriptions, I also have one lying here, and this is the new medicine I got, and I don't understand it, because I should have asked about it.

W3: Yes, I have these blue folders, you probably don't know them, but those blue boxes for morning, midday, evening, and there is for (cough)...think there is for eight days, probably, that can be right? Eight days, I believe that, and I sort them every second week. I sit by the dining table, and line the whole thing up, and then I sort them.

The aforementioned examples depict 2 different scenarios of handling and understanding medicine. In the first scenario, one of the informants blames herself for not seeking information about the new medication when she first started taking it because she is unfamiliar with it. While she accepts responsibility for her medication, she still requires assistance in understanding it. By contrast, the second informant has established routines for managing her medication, and therefore, understands what she deposits into the pillbox.

As a result of errors and inconsistent care from municipal employees, the informants began to question their trust in the municipality's care team. Although the informants could receive assistance from the municipality with their prescriptions and medication management, they found that the assistance and knowledge provided by municipal employees lacked the necessary qualifications.

W2: ...but she wasn't, and then she made the mistake of repeating after the other, and I quickly notice it, and it's not, it is not calcium tablets we sit and play with.

This is a serious concern, as indicated by the informant's statement that the pills are not calcium supplements, and taking medication in incorrect amounts could have negative effects on her health. Therefore, it is vital that her medication is prescribed correctly for her condition.

When it comes to medication, there was a strong tendency among the informants to rely on the RCC nurses, especially during exacerbations. The RCC nurses use telephone communication to reach out to the informants and inquire about their health. If necessary, the nurses may advise the informants to take additional medication. The informants comply with the

XSL•FO
RenderX

nursing advice and adjust their medication accordingly because they have a high level of trust in the digital nurses.

> *M2: Yes, "Nærklinikken" is the ones who change it now, yes, they just say you have to take two breaths in the morning, and they do it regularly if I have felt worse for a little while. I'll just get more, double up.*
>
> *M2: Yes, I feel very safe.*

When it comes to medication, the informants trust the guidance of the digital nurses because they feel it is their responsibility to adjust their medication. They feel secure knowing that others are assisting them and providing direction with their medication management.

When the RNs oversee the health condition of the informants, their independence in managing their chronic disease and their understanding of their medication do not seem to improve. It appears that the RNs are still somewhat paternalistic. However, the informants do experience a sense of safety, particularly when it comes to their health and medication.

Value Creation

Behavior

This category identifies the former daily routines that had to change or be excluded from the informants' lives because of COPD. Furthermore, it highlights the new routines that should be adapted because of the weakened ability caused by their condition.

The informants must adapt their daily routines to accommodate their diminished capacity because of COPD, necessitating the establishment of new habits. Consequently, they may take fewer walks, experience reduced appetite, or sleep longer than usual. This limitation often confines the informants to their residences. When queried about their daily lives, the informants provided a range of responses. Some spoke very briefly and exhibited a negative attitude, while others believed that obtaining a comprehensive understanding of how COPD impacts daily activities was crucial.

> *Interviewer: ...oh if you should tell me how a typical everyday looks for you K1, what do you do on a general day?*
>
> *W1: Sitting here*
>
> *Interview: You sit there*
>
> *W1: Yes (cough), but sometimes when I'm well, I go out in the garden.*

In this case, the informant primarily spends time sitting on the same couch and does not elaborate much on her everyday activities. She finds joy in moving outside and into the garden whenever possible. Previously, she engaged in various artistic activities and housework as part of her daily routine, but these tasks are no longer feasible due to her health.

Some informants expressed that it was still important to maintain cleanliness in their own homes. While the municipality provides cleaning assistance to the majority of the informants, some individuals still prefer to handle specific tasks on their own.

> *W2: ...I said to her, now don't think I'm crazy but I've been standing and ironed my bedsheets for several days, then she was about to faint (...). Only the elderly irons their bedsheets. I have always done it, and I will not stop doing it as long as I can stand on my feet.*

She continues to prioritize tasks such as making her bed and changing the sheets, as she has always done. Despite the challenges posed by her health, she decides to persist because these tasks hold significant importance for her. However, she acknowledges that it may take several days to complete them. By contrast, most of the informants expressed overall dissatisfaction with the cleaning assistance provided by the municipality, stating that they often had to make numerous corrections.

Another aspect they felt had changed because of their condition was the rhythm of the day. They found that getting out of bed in the morning was becoming more challenging, or they noticed that they were waking up earlier. This change could be attributed to their increased frequency of sleeping and reduced engagement in everyday activities.

> *M1: Yes. Well, but I wake up before Satan gets his shoes on, because I am used to doing something, and I cannot really more, so I never get really really tired (coughing), so I do not get so terrible many hours of sleep (coughing).*

Here it is highlighted that the informant's daily rhythm has shifted from its previous pattern, and the indication is that their lack of sleep stems from both reduced activity and diminished tiredness. None of the informants mentioned experiencing anxiety or shortness of breath during the night, which could also contribute to a different daytime rhythm. However, the increased need for sleep during the day was frustrating, as it could result in missing out on certain activities.

Hobbies

The informants engaged in different activities in their lives that held personal value for them, and some had to alter their activities because of COPD. This category focuses on the activities that the informants currently undertake and have previously engaged in.

The activities and interests of the informants varied depending on their weekly or daily routines. However, their condition often took precedence over their interests, and the activities that were feasible differed among the informants. Additionally, there was a gender disparity, with men favoring fishing and other outdoor activities, while women tended to engage in activities such as handicrafts.

> *M2: So now that I have been sick, yes, I go out and fish a little, then I go and help a little with some horses.*
>
> *W3: ...and then I'm knitting or doing the crosswords or trying to sew on the sewing machine (laughing).*
>
> *M1: yeah, I'm trying to do the things I care about and like, unfortunately I can't paint anymore, as I cannot stand the smell of turpentine anymore, it's sad*

XSL•FO
RenderX

because I've spent a lot of time painting, I don't have the energy to start writing more books. So I, I read a little, it's a bit difficult now with these glasses, but I've read a lot, and I get a lot of pleasure from it.

The different accounts provided by the informants offer insights into how interests are possible for individuals with chronic illness as well as how the condition can prevent them from pursuing activities they like/enjoy. Despite certain interests being curtailed by the condition, informants still strive to engage in activities that bring them joy and hold value for them. However, in some cases, informants found it challenging to pursue their interests because of the awareness of potential shortness of breath.

The informants expressed similar daily desires and willingness to go outside, but their walks had become shorter over time. Occasionally, they cited the weather as an excuse to stay indoors.

W2: Then the little dog and I go in and rest for an hour or half an hour, and lull a little and sleep a little. Perhaps it is something completely different. Then I get up and get ready, and then I go with my little dog and pick up the newspaper in the mailbox. We used to go for longer walks, but I don't unfortunately, I can't do it anymore.

W3: For just such a trip, so there is not much nature to go and look at from here and down to the municipal office, but just to get out and get some fresh air

The informant acknowledges here that the challenge of going on longer walks is something that annoys her, but she simply cannot manage it anymore due to shortness of breath. Despite this limitation, going for a walk can bring relief to the informants. Even though nature may not always be visually stunning, the informant finds solace in being able to get outside, especially on slightly gloomier days.

Motivation

This category underscores how the informants experienced a lack of motivation to engage in daily activities. Throughout the conversations, there was a tendency to discuss things they would love to do but lacked the competence or strength to accomplish, or they invented excuses because the tasks seemed overwhelming. These could range from simple tasks such as planting a rose to more complex endeavors such as writing a book, attending gym classes, or cooking.

W1: Not at all, and then you lose the motivation.

Interviewer: Yes, I can understand, if you have been somewhere where you think it was good, and then you come to something else that you don't think is at the same standard...

W1: but then, okay, I am not...there, but oh its hunting me, when I have to go. I will probably just get it over with, right?

The informant's desire and motivation to participate in a COPD exercise team depend on how the teacher conducts the sessions. A negative experience with teaching methods in the past has diminished the informant's motivation, making it difficult for them to participate. It has transformed from an activity that brings joy to feeling like an obligation, something that the informant feels they must do rather than something they want to do.

Resources

Social Networks

This category explores the social networks that the informants are a part of and how they use them in their day-to-day activities.

The size of the social network varied among the informants, but the significance of social interactions was equally important to all of them. Family relationships showed considerable diversity among the informants, with some maintaining close contact with their family through daily conversations, weekly scheduled visits, or having their spouse present. By contrast, there were some informants who had limited contact with their families and spent much of their daily lives alone.

W2: I miss him very much. We lived in Fyn, and we talked over the phone several times a week. I miss my family very much, and I also miss my friends. And they all passed away...

W2: ...I talk to my daughter. So so I don't talk about..., she can say to me, I think you sound a little stalled mother, because then I'm just for the moment and then, and then we're not talking about it anymore. She knows what it is, but we don't need to.

The above description indicates that the informant is alone due to deaths in the family and social circles, and there are a significant number of people missing from her family. Interestingly, the informant does not mention at the present time that she still has 2 daughters, which she only brings up later in the interview. The informant's description of the varying family relations indicates that she is left more isolated and alone, which telephone conversations with her daughters cannot fully mitigate. The conversations surrounding the family and friends of the "lonely" informants were marked by a sense of sadness and depression over the lack of contact.

Informants with close family relations expressed how their family and close relationships maintain continuous contact with them. The conversations were even interrupted by phone calls from their families, highlighting the frequent and ongoing nature of their communication.

M2: Yeah, she is calling, or she has stopped a little, but otherwise she calls every morning, around 9 o'clock or something like, "How do we breathe today?" She says then, (laughing)

W1: Then I also have my granddaughter, I talk a lot with her, but I also take care of what I said to her, because she is a little unstable.

There are 2 different scenarios for contact described here, both indicating that the informants have contact with their families, signifying close relationships. In one instance, it is the informant's mother showing interest in their self-monitoring and health status. In the second scenario, the informant not only maintains close contact, but also plays a protective role for her grandchild, who also suffers from a diagnosis. Despite varying

XSL•FO
RenderX

family dynamics, it is evident that the informants can be divided into 2 groups: those with close family relations and those lacking such connections.

Their interest in engaging in social activities was also significant, but the informants often found themselves coming up with excuses for not participating or found it challenging to leave their homes.

> *W3: Yes, I haven't reached it yet, but I'm probably getting enough. There has been something on Friday, because otherwise I had set myself up for, I have otherwise gone to gym down in Vig, but.. that, which is quite far from the station off and down to Balsagård (...), of course I can go down there from time to time, but as I have it, oh for the last season there, I was not there quite many times, but it costs no matter what, they do not pull anything from because I have not been there, and that.. it annoys me a little. Then there is Red Cross that has something like this in high town.. exercise, sport is known enough, and it is every Friday morning, but there has been something here the last couple of Fridays, and I also have to just get into the rhythm that I have to go there until half 10 p.m.*

As emphasized, the informant highlights that traveling a long distance to attend gymnastics is a major obstacle. Despite continuing to pay for it, this does not motivate her to attend regularly. She also mentions the challenge of incorporating it into her Friday routine and making it a regular part of her schedule.

They did not envision themselves participating in social events related to their COPD. Some of them were members of Lungeforeningen, the Danish Association for Lung Diseases. Although the Lung Association organizes various gatherings for those with COPD, none of them appeared interested in attending.

Health Professionals

The category focuses on the informants' interaction with various health professionals and how those relationships hold significance for them.

The relationship with health professionals was highly significant, as it was essential for the informants to feel secure while also being with mutual respect and seriousness. The informants interacted with various health professionals in their daily lives, and this analysis distinguished the difference between "ordinary" health professionals and PreCare nurses.

The informant's relationship with the assistance offered is crucial.

> *W2: After a hospitalization for yeah I don't know, let's just say a year ago. There seems to be, I can't remember who thinks that there should be a home care and dosing the medicine. And now you must not misunderstand me because I am not a racist, but there comes a little colored girl who could not really speak Danish and she was not very sweet if she had been sweet and smiling, pleasant, then it would have been something else, but it was she not, and then she made mistakes twice after the other, and I discovered it quickly, and it is not, it isn't the tablets we sit and play with.*

According to the informant, the connection with home care has been challenging because of mistakes and uncertainty. As a result of this, the informant has lost faith in the home care, which should be there to lend a helping hand and not cause her problems on a regular basis.

Unlike other health care professionals, the nurses in PreCare have succeeded in establishing a sense of security and mutual respect with the informants. The collaboration with those involved in their COPD care instills confidence in the PreCare project and the nurses among the informants.

> *M1: There is most of the contact through the nurses, just to start, just when I started up there was a doctor who was here, and so I have nothing bad to say about him, oh and it is also those who prescribe some medicine if I lack it, and such some things not too (slang). I think I have a good relationship with them, and are really pleased to have them, oh...and feel there is a great confidence to be with them. So, as I said earlier, I was sure I would have become a burden for the hospital if I had not known them. The society saves money, and that's not bad.*

> *W2: But Nærklinikken has helped me, exceptionally. I'm glad I got in touch with you, you can believe. I don't know what I would have done without medical care. They do nothing.*

The informant expresses happiness for the nurses and thinks that their connection is good. The following description includes several elements though. The informant mentions that their participation in PreCare makes them feel like less of a burden for the hospital and the municipality, which holds significant meaning for them. He also expresses faith in the nurses, which he had expressed several times in the conversation. Being a part of a project that highlights the superior care provided by nurses compared with general medical care has been particularly significant for the female informant.

Nurses are not only available to informants, but also offer support if an informant's condition deteriorates. When an exacerbation happens, trust means the informant has no reason to question the nurses and takes the prescribed prescription without a second thought.

> *M2: Yes, Nærklinikken says it, it is the ones who change it now, yes, they say you just have to have two breaths in the morning, and they do it regularly if I have had it a little bad for a while. No, then I'll just be put up, double up.*

Without hesitation, the informant chooses to follow the nurses' recommendations. He has completely surrendered to the project, giving them full responsibility for his condition.

Technology

This category covers the informants' daily technology and their search for health-related information.

XSL•FO
RenderX

All informants admitted to having technology at home, although the way they used it and its purpose varied. For some, technology provides entertainment during moments of boredom in their daily lives.

> *W1: Yes, I'm mostly on the computer when I get bored.*
>
> *Interviewer: What are you doing on the computer?*
>
> *W1: I am playing games*
>
> *W1: I have two different games I have discovered.*

Technology was not utilized by the informants as a means to gather information about COPD. The informants felt they already possessed all the information they needed about COPD, and they were concerned that obtaining more information might increase their anxiety.

> *M1: It is very very rare; it is very rare. If I happen to hear that there is something new about it, then I can well find out to look it up, that it is not so exciting to read about, so*
>
> *W1: I think the more you read, the more nervous you become.*

The informants do not use technology to seek information about their condition. The informants did not mention being part of online groups where information could be shared during the interviews. None of the informants mentioned using social media platforms such as Facebook as a community for sharing information about their condition. They also emphasized that they generally did not share information about their condition through digital solutions.

> *M3: It irritates me sometimes when we sit, sometimes I cut through and say, now we don't want to talk about illness, because, oh, then such a short evening can go*

The medical equipment provided by PreCare did not pose any problem for the informants to use on a daily basis. They all expressed how easy it is to use it and how it takes only a few seconds to use the technology. They appeared confident and stated that they performed the measurements every day.

> *M4: it's so easy, that's in order...the only thing is now just, the crazy computer goes out, or (...), I can't restart, even though I have PIN code...no matter what I do, it won't, so I wait when it comes a past...so it can restart again, the only problem...*

Despite the ease, they experienced some issues with the devices. The informants encountered issues with logging in, forgot to charge the tablet, or even misused the thermometer.

Discussion

Principal Findings

This study offers valuable insights into understanding the needs, values, and preferences of individuals living with COPD as well as which communities they identify with in a digital context. Indeed, the findings highlight that while fluctuations in their health condition significantly affect the daily lives of the informants, factors such as having hobbies, old habits, and social connections play a crucial role in their overall well-being. This underscores the importance of recognizing individuals with COPD as complete human beings beyond their medical condition.

It is interesting to note the distinction between the 2 communities the informants belong to. The community centered around the RCC represents a vital support network for them, where they feel included and have developed trust with the staff. This highlights the importance of such digital health services in providing continuous support and guidance for individuals managing chronic conditions such as COPD. Involving close relatives in the community centered around the RCC can further enhance the support system for individuals with COPD. The other community involves participation in social activities outside the context of their condition which provides informants with opportunities for social interaction, enjoyment, and connection with others beyond their health concerns. These activities offer a sense of normalcy and contribute to their overall quality of life, allowing them to engage in meaningful relationships and experiences beyond the realm of COPD management. Even though the informants may face constraints due to their condition, they still find value in participating in social activities, even if their involvement is limited.

The Needs, Values, and Preferences in the Digital Context

The informant's emphasis on maintaining their daily activities underscores the significance of preserving their sense of normalcy and independence despite their COPD symptoms. It reflects their desire to continue living fulfilling lives and not be defined solely by their health condition. Symptoms often contribute to a lower quality of life and well-being. The findings from an earlier study [32] resonate with the experiences reported by individuals with COPD in this study. Breathlessness, a common symptom of COPD, can significantly impact an individual's quality of life by limiting their ability to engage in daily activities and causing distress. This aligns with the participants' reports of reduced quality of life related to breathlessness, highlighting the importance of addressing this symptom to improve overall well-being for individuals living with COPD. While outdoor activities may be affected by COPD, the focus of the informants seems to be more on how the disease impacts their ability to engage in everyday tasks and maintain their hobbies or household chores. This also resonates well with another study including interviews of patients with COPD [7]. The study revealed that for women, being active in housekeeping was important and valued, while for men, maintaining the garden held similar significance. They also reported that people with COPD prefer activities within their home or immediate vicinity, showing little interest in engaging in social activities located far away due to their reduced physical abilities. The need to be close to a safe environment was similar for some of the informants in this study. However, some reported that they did leave their house. For example, one informant was helping with some horses, while others needed to take their dog to the dog groomer. Access to virtual support may play an important role here, as the informant could call the RN in the RCC at any time if anything were to happen.

XSL•FO
RenderX

The COPD-Related Conditions' Impact on the Identity

For the informants, it was important to avoid discussing or being associated with their COPD condition by shifting the topic during the interviews. This aligns with a previous study where people with COPD were not interested in being identified solely by their illness [33]. They found that their identity related to their condition, termed "illness identity," was affecting their roles and could potentially separate them from the social network or community. An additional contributing factor to social segregation was a sense of having a self-inflicted disease, leading to feelings of shame and guilt [33]. However, this was not evident in our data.

Most, but not all, of the informants appear to be able to cope with their diagnosis and condition. They try to continue the same kind of interactions and activities while considering their condition's restrictions. As a result, individuals experience a sense of maintaining their own identity within their communities, yet occasionally feel the sense of lacking something. This can involve engaging in distant activities, such as joining a choir.

The Two Communities of Practice

The Community of COPD Practice

In relation to the community with its formal caregivers, the informants experienced a genuine interest from the RNs in their well-being and they provided them with support in an empathic way. This experience may be attributed to the PreCare environment, with free access to RNs 24/7, where they always kept an eye on the informants. This may explain the absence of "anxiety" in the interviews with the informants. These could be attributed to the RN's *ability* to provide immediate support in response to changes in their health condition, with medication for deterioration accessible at home [32]. Anxiety, which often dominates the daily lives of individuals with COPD, is thus better managed [32]. Therefore, it is necessary for the RNs to be trained to instill confidence in individuals with COPD or other LTHCs, enabling them to feel more independent and socially active with the RCC, their equipment, and a medicine box readily available, thereby reducing anxiety and maximizing the benefits of their resources.

The informants' immediate access to the RNs in the RCC appears crucial, as it enhances their self-efficacy and confidence in how their equipment aids them in managing deterioration.

However, despite feeling secure in their use of the equipment, the informants did not appear to be influenced by their ability for self-management and did not feel more empowered, likely due to experiencing the RNs as being paternalistic. The informants were unable to fully benefit from the virtual support environment due to the influence of the RNs and instead remained in a passive role.

RNs and other clinicians will need to be aware of how they communicate to facilitate a dialog that is not experienced as paternalistic, but rather as a coaching conversation.

The Community of Social Practice

In relation to their social communities, the extent of social relationships varied among the informants. The importance of a family community aligns with the findings of Nicolson and Anderson [32], who showed that family and relatives significantly influence the quality of life for people with COPD. Nicolson and Anderson [32] identified that COPD impacted how individuals connected with relatives and perceived their ability to fulfill their roles within the family. In contrast to the study by Nicolson and Anderson [32], our findings indicate that the informants' family roles were not influenced by their COPD. They maintained their roles and continued normal interaction with their relatives and families.

Not all informants participated in social activities or were part of a local community. Those without support from friends and families experienced difficulties engaging in and finding motivation for social activities. They felt lonely when left alone in their homes, whereas those with a social network experienced loneliness to a lesser degree. This contrasts with another study [7], which found that loneliness was also a major issue for those with family support, such as spouses and friends [7]. Those who felt lonely because of their lack of participation in social activities found some comfort in the availability and contact with the RN, which to a certain extent reduced their experience of loneliness.

This underscores the necessity for RNs to possess skills in mental and social support, which should be included in the education programs for nurses.

Online resources such as patient portals and social media (eg, Facebook) can constitute a community for people with COPD. However, despite the availability of these platforms to our informants and their daily use of tablets, none of them considered these online opportunities in relation to their COPD condition. This may be due to various reasons. Some informants felt they had sufficient knowledge or were unsure how to interpret the overwhelming information on the internet. Additionally, they may have wanted to avoid exposing their diagnosis or involving others outside their close network [34]. The role of the PreCare environment and the RNs may substitute the need for a social media platform or COPD-related conversations on platforms such as Facebook.

When the informants do not participate in online communities, they may miss the opportunity to access new knowledge or learn from others with similar conditions. This lack of engagement can reduce their ability to manage their condition effectively and hinder their empowerment. Joining an online community and being actively involved can help transform newcomers into "super users" and "experts" [35]. These "online experts" can then help other members of the community, forming a virtual community of practice [25]. Thus, participants in the PreCare project may miss the opportunity to develop into experts through online activities but may instead develop this competence through participation in other communities or collaboration with the RNs.

XSL•FO
RenderX

Limitations

The study is based on 7 informants. This may be considered a limitation, as the relatively small number may result in some perspectives of people living with COPD not being expressed in the data. However, as all informants are exposed to the same PreCare environment, have the same diagnosis of COPD, and live within the same area, we find that the necessary number of participants to have enough power of information is met [36]. This is supported by the presence of common patterns among the informants and the alignment of the overall findings with the data obtained in the PreCare project. Further studies are needed to confirm our findings before they can be considered valid for scaling up and evaluating the impact of working with landscapes or communities of practice. This support aims to foster a sense of more active and independent living based on existing values.

Perspective

The findings suggest that education for RNs and other health professionals should focus on their roles as professionals while also acting as facilitators. They should avoid being paternalistic to create a space for the development of self-efficacy and self-management. A motivating factor will help develop self-efficacy and confidence, enabling people with COPD to be more socially active and encouraging them to pursue their desires. Health professionals play a key role here, as they can provide the means to help individuals become more active, thereby increasing their well-being.

Conclusions

When using digital health solutions, people's needs, values, and preferences should be considered, focusing primarily on addressing the whole person rather than just the "illness." This approach creates the best opportunity for individuals to maintain their daily activities and feel empowered. The 2 communities the informants take part must work together and will intersect in their daily lives. They should support each other, involving the needs, values, and preferences of the individuals, and ensuring that upcoming digital health services include and embrace situated learning to enhance people's empowerment and self-management. Furthermore, new educational programs should be developed or considered to enhance the competencies of RNs who are involved in digital health services. This will provide the best opportunity for the 2 communities to collaborate and support the daily activities of people with chronic conditions.

Acknowledgments

The authors thank the PreCare Project for assisting with the recruitment of informants and the informants for participating in the study. The first author has received funding from Region Zealand. CWS, KB, and LK have been partly financed by the EU H2020 Research & Innovation Programme (grant agreement number 101016848) for the period 2021-2024.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Interview guide.
[DOCX File , 16 KB - humanfactors_v11i1e53131_app1.docx]

Multimedia Appendix 2
Categorization of the data.
[XLSX File (Microsoft Excel File), 12 KB - humanfactors_v11i1e53131_app2.xlsx]

References

1. Kayser L, Phanareth K. Leave no one behind: the role of digital health literacy. European Respiratory Society 2023:79-94. [doi: 10.1183/2312508x.10001023]
2. Cavanagh R, Gerson SM, Gleason A, Mackey R, Ciulla R. Competencies Needed for Behavioral Health Professionals to Integrate Digital Health Technologies into Clinical Care: a Rapid Review. J Technol Behav Sci 2022 Feb 17;8(4):1-14 [FREE Full text] [doi: 10.1007/s41347-022-00242-w] [Medline: 35194561]
3. Jarva E, Oikarinen A, Andersson J, Tuomikoski A, Kääriäinen M, Meriläinen M, et al. Healthcare professionals' perceptions of digital health competence: A qualitative descriptive study. Nursing Open 2022 Jan 30;9(2):1379-1393. [doi: 10.1002/nop2.1184]
4. Longhini J, Rossettini G, Palese A. Digital Health Competencies Among Health Care Professionals: Systematic Review. J Med Internet Res 2022 Aug 18;24(8):e36414. [doi: 10.2196/36414]
5. Chronic obstructive pulmonary disease (COPD). Chronic obstructive pulmonary disease (COPD). URL: https://www.who.int/news-room/fact-sheets/detail/chronic-obstructive-pulmonary-disease-(copd) [accessed 2023-01-30]
6. Nøgletal OL. Nøgletal om lungesygdomme. Lungeforeningen. 2016. URL: https://tinyurl.com/5dsabf2s [accessed 2023-01-30]
7. Williams V, Bruton A, Ellis-Hill C, McPherson K. What really matters to patients living with chronic obstructive pulmonary disease? An exploratory study. Chron Respir Dis 2016 Jun 29;4(2):77-85. [doi: 10.1177/1479972307078482]

XSL•FO
RenderX

8. Global Initiative for Chronic Obstructive Lung Disease - GOLD. GOLD Report 2023. URL: https://goldcopd.org/2023-gold-report-2/ [accessed 2023-05-25]
9. GOLD. Global strategy for the diagnosis, management, and preventien of chronic obstructive pulmonary disease (2018 report). GOLD. 2018. URL: https://tinyurl.com/3r6mkuuk [accessed 2023-05-25]
10. Pare G, Jaana M, Sicotte C. Systematic Review of Home Telemonitoring for Chronic Diseases: The Evidence Base. Journal of the American Medical Informatics Association 2007 May 01;14(3):269-277. [doi: 10.1197/jamia.m2270]
11. McDowell JE, McClean S, FitzGibbon F, Tate S. A randomised clinical trial of the effectiveness of home-based health care with telemonitoring in patients with COPD. J Telemed Telecare 2015 Mar;21(2):80-87. [doi: 10.1177/1357633X14566575] [Medline: 25586812]
12. Worth H, Dhein Y. Does patient education modify behaviour in the management of COPD? Patient Education and Counseling 2004 Mar;52(3):267-270. [doi: 10.1016/s0738-3991(03)00101-0]
13. Nissen L, Lindhardt T. A qualitative study of COPD-patients' experience of a telemedicine intervention. International Journal of Medical Informatics 2017 Nov;107:11-17. [doi: 10.1016/j.ijmedinf.2017.08.004]
14. McCabe C, McCann M, Brady AM. Computer and mobile technology interventions for self-management in chronic obstructive pulmonary disease. Cochrane Database Syst Rev 2017 Dec 23;5:CD011425. [doi: 10.1002/14651858.CD011425.pub2] [Medline: 28535331]
15. Hardman R, Begg S, Spelten E. What impact do chronic disease self-management support interventions have on health inequity gaps related to socioeconomic status: a systematic review. BMC Health Serv Res 2020 Feb 27;20(1):150 [FREE Full text] [doi: 10.1186/s12913-020-5010-4] [Medline: 32106889]
16. WHO. Social determinants of health. URL: https://www.who.int/health-topics/social-determinants-of-health [accessed 2023-09-14]
17. Madsen J, Kanstrup AM, Josephsson S. The assumed relation between occupation and inequality in health. Scand J Occup Ther 2016 Aug 28;23(1):1-12. [doi: 10.3109/11038128.2015.1075065] [Medline: 26314738]
18. Lave J, Wenger E. Situated Learning. Cambridge, United Kingdom: Cambridge University Press; 1991.
19. Bandura A. Social Learning Theory. Canadian Journal of Sociology / Cahiers canadiens de sociologie 1977;2(3):321. [doi: 10.2307/3340496]
20. Spradley JP. Participant Observation. Anthropological Quarterly 1980 Oct;53(4):260. [doi: 10.2307/3318111]
21. WHO. WHO Framework on integrated people-centred health services. 2016 Aug. URL: http://www.who.int/servicedeliverysafety/areas/people-centred-care/en/ [accessed 2023-09-14]
22. Phanareth K, Vingtoft S, Christensen AS, Nielsen JS, Svenstrup J, Berntsen GKR, et al. The Epital Care Model: A New Person-Centered Model of Technology-Enabled Integrated Care for People With Long Term Conditions. JMIR Res Protoc 2017 Jan 16;6(1):e6 [FREE Full text] [doi: 10.2196/resprot.6506] [Medline: 28093379]
23. Data Development, Region Zealand. PRECARE: ViosKlinikken - Sector Neutral, Data Driven and Person-Centered Health and Care Services. PreCare Clinic 2024:140.
24. Kvale S. Interview, en introduktion til det kvalitative forskningsinterview. In: Interview, en introduktion til det kvalitative forskningsinterview. Denmark: Hans Reitzel Forlag; 1990.
25. Wenger E. Communities of Practice. Cambridge: Cambridge University Press; 1999.
26. Kayser L, Rossen S, Karnoe A, Elsworth G, Vibe-Petersen J, Christensen JF, et al. Development of the Multidimensional Readiness and Enablement Index for Health Technology (READHY) Tool to Measure Individuals' Health Technology Readiness: Initial Testing in a Cancer Rehabilitation Setting. J Med Internet Res 2019 Feb 12;21(2):e10377 [FREE Full text] [doi: 10.2196/10377] [Medline: 30747717]
27. Terp R, Kayser L, Lindhardt T. Older Patients' Competence, Preferences, and Attitudes Toward Digital Technology Use: Explorative Study. JMIR Hum Factors 2021 May 14;8(2):e27005 [FREE Full text] [doi: 10.2196/27005] [Medline: 33988512]
28. Terp R, Lindhardt T, Kayser L. Theory-driven development of an educative nutritional intervention (ENI) supporting older hospital patients to eat sufficiently, assisted by an eHealth solution: an intervention mapping approach. BMC Health Serv Res 2022 Nov 28;22(1):1435 [FREE Full text] [doi: 10.1186/s12913-022-08679-8] [Medline: 36443804]
29. Elo S, Kyngäs H. The qualitative content analysis process. J Adv Nurs 2008 Apr;62(1):107-115. [doi: 10.1111/j.1365-2648.2007.04569.x] [Medline: 18352969]
30. Nvivo. Nvivo. Alfasoft. URL: https://help-nv.qsrinternational.com/12/win/v12.1.115-d3ea61/Content/welcome.htm [accessed 2021-01-30]
31. The Danish National Center for Ethics. Overview of Mandatory Reporting. 2024 Feb 01. URL: https://researchethics.dk/information-for-researchers/overview-of-mandatory-reporting [accessed 2024-03-25]
32. Nicolson P, Anderson P. Quality of life, distress and self-esteem: a focus group study of people with chronic bronchitis. Br J Health Psychol 2003 Sep 16;8(Pt 3):251-270. [doi: 10.1348/135910703322370842] [Medline: 14606972]
33. Lippiett K, Richardson A, May CR. How do illness identity, patient workload and agentic capacity interact to shape patient and caregiver experience? Comparative analysis of lung cancer and chronic obstructive pulmonary disease. Health Soc Care Community 2022 Nov 28;30(6):e4545-e4555 [FREE Full text] [doi: 10.1111/hsc.13858] [Medline: 35633149]
34. Andreou A, Dhand A, Vassilev I, Griffiths C, Panzarasa P, De Simoni A. Understanding Online and Offline Social Networks in Illness Management of Older Patients With Asthma and Chronic Obstructive Pulmonary Disease: Mixed Methods Study

XSL•FO
RenderX

Using Quantitative Social Network Assessment and Qualitative Analysis. JMIR Form Res 2022 May 17;6(5):e35244 [FREE Full text] [doi: 10.2196/35244] [Medline: 35579933]
35. Joglekar S, Sastry N, Coulson NS, Taylor SJ, Patel A, Duschinsky R, et al. How Online Communities of People With Long-Term Conditions Function and Evolve: Network Analysis of the Structure and Dynamics of the Asthma UK and British Lung Foundation Online Communities. J Med Internet Res 2018 Jul 11;20(7):e238 [FREE Full text] [doi: 10.2196/jmir.9952] [Medline: 29997105]
36. Malterud K, Siersma VD, Guassora AD. Sample Size in Qualitative Interview Studies: Guided by Information Power. Qual Health Res 2016 Nov 10;26(13):1753-1760. [doi: 10.1177/1049732315617444] [Medline: 26613970]

Abbreviations

COPD: chronic obstructive pulmonary disease
ECM: Epital Care Model
GOLD: The Global Initiative for Chronic Obstructive Lung Disease
LTHC: long-term health condition
RCC: response and coordination center
RN: registered nurse

Edited by A Choudhury; submitted 27.09.23; peer-reviewed by H Salim, J Turnbull; comments to author 31.12.23; revised version received 29.04.24; accepted 24.05.24; published 21.06.24.

Please cite as:
Schmidt CW, Borgnakke K, Frølich A, Kayser L
Preferences, Needs, and Values of Patients With Chronic Obstructive Pulmonary Disease Attending a Telehealth Service: Qualitative Interview Study
JMIR Hum Factors 2024;11:e53131
URL: https://humanfactors.jmir.org/2024/1/e53131
doi:10.2196/53131
PMID:

©Camilla Wong Schmidt, Karen Borgnakke, Anne Frølich, Lars Kayser. Originally published in JMIR Human Factors (https://humanfactors.jmir.org), 21.06.2024. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on https://humanfactors.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Predictive Factors of Physicians' Satisfaction and Quality of Work Under Teleconsultation Conditions: Structural Equation Analysis

Liliana Hawrysz[1*], PhD, DSC; Magdalena Kludacz-Alessandri[2*], PhD, DSC; Renata Walczak[3*], PhD, DSC

[1]Faculty of Management, Wrocław University of Science and Technology, Wroclaw, Poland
[2]College of Economics and Social Sciences, Warsaw University of Technology, Plock, Poland
[3]Faculty of Civil Engineering, Mechanics and Petrochemistry, Warsaw University of Technology, Plock, Poland
[*]all authors contributed equally

Corresponding Author:
Renata Walczak, PhD, DSC
Faculty of Civil Engineering, Mechanics and Petrochemistry
Warsaw University of Technology
17 Lukasiewicza St
Plock, 09-400
Poland
Phone: 48 887780100
Email: renata.walczak@pw.edu.pl

Abstract

Background: The COVID-19 pandemic contributed to an increase in teleconsultation adoption in the Polish primary health care system. It is expected that in the long run, teleconsultations will successfully replace a significant part of face-to-face visits. Therefore, a significant challenge facing primary health care facilities (PHCs) is the acceptance of teleconsultations by their users, especially physicians.

Objective: This study aimed to explore physicians' acceptance of teleconsultations during the COVID-19 pandemic in Poland.

Methods: A representative survey was conducted among 361 physicians of PHCs across Poland in 2021. For the purposes of the study, we developed a modified Technology Acceptance Model (TAM) model. Based on the modified TAM, we analyzed the impact of perceived usefulness (PU), perceived ease of use (PEU), and intention to use teleconsultation (INT) on physicians' satisfaction (SAT) and quality of work (Q). The psychometric properties of the research instrument were examined using exploratory factor analysis. Finally, structural equation modeling was used for data analysis.

Results: The results indicated a generally high level of PU (mean 3.85-4.36, SD 0.87-1.18), PEU (mean 3.81-4.60, SD 0.60-1.42), INT (mean 3.87-4.22, SD 0.89-1.12), and SAT (mean 3.55-4.13, SD 0.88-1.16); the lowest rated dimension in TAM was Q (mean 3.28-3.73, SD 1.06-1.26). The most important independent variable was PU. The influence of PU on INT (estimate=0.63, critical ratio [CR]=15.84, $P<.001$) and of PU on SAT (estimate=0.44, CR= 9.53, $P<.001$) was strong. INT was also a key factor influencing SAT (estimate=0.4, CR=8.57, $P<.001$). A weaker relationship was noted in the effect of PEU on INT (estimate=0.17, CR=4.31, $P<.001$). In turn, Q was positively influenced by INT (estimate=0.179, CR=3.64, $P<.001$), PU (estimate=0.246, CR=4.79, $P<.001$), PEU (estimate=0.18, CR=4.93, $P<.001$), and SAT (estimate=0.357, CR=6.97, $P<.001$). All paths between the constructs (PU, PEU, INT, SAT, and Q) were statistically significant, which highlights the multifaceted nature of the adoption of teleconsultations among physicians.

Conclusions: Our findings provide strong empirical support for the hypothesized relationships in TAM. The findings suggest that the PU and PEU of teleconsultation have a significant impact on the intention of physicians to adopt teleconsultation. This results in an improvement in the satisfaction of Polish physicians with the use of teleconsultation and an increase in Q. The study contributes to both theory and practice by identifying important prognostic factors affecting physicians' acceptance of teleconsultation systems.

(JMIR Hum Factors 2024;11:e47810) doi:10.2196/47810

KEYWORDS

perceived ease of use; perceived usefulness; physicians' satisfaction; behavioral intention to use telemedicine; health care quality; technology acceptance model; TAM; COVID-19: telemedicine

XSL•FO
RenderX

Introduction

Background

Telemedicine is an IT-based method that has the potential to support and enhance physicians' patient care [1]. Telemedicine is defined as a tool using information and communication technology (ICT) that is used to support and promote remote care, health-related professional education, and public health administration [2]. One of the basic forms of telemedicine is teleconsultations. Medical consultations have been provided remotely (teleconsultations), by telephone, or by instant messaging for many years, but the COVID-19 pandemic has forced widespread use of this form of communication with patients. Teleconsultations have reduced the high costs of medical services and the queues of patients in clinics [3]. In Poland, the need for teleconsultations began at the beginning of the pandemic. Most physicians had to adapt to the requirements of COVID-19 rules and regulations. Even though the need for teleconsultations has already been officially abolished, telemedicine is still the future of medicine. Not all advice has to be given to patients on-site in a clinic. In the face of a decreasing number of physicians, an aging society, and an increased demand for health care, telemedicine seems to be a solution that will solve the problems of personnel shortage. However, the use of telemedicine tools requires the medical staff to be proficient in using them and to accept and support such solutions.

The aim of our study was to examine the physicians' satisfaction (SAT) and quality of work (Q) in conditions of teleconsultations during the COVID-19 pandemic in Poland. For the purposes of the study, we developed a modified Technology Acceptance Model (TAM) and tested it in primary health care facilities (PHCs) in Poland. This model allowed us to analyze the impact of perceived usefulness (PU), perceived ease of use (PEU), and intention to use teleconsultation (INT) on SAT and Q. In our study, we focused primarily on teleconsultations, consisting of telephone and video conversations between physicians and patients [2]. Teleconsultations are the basic form of telemedicine in Poland, and during the COVID-19 pandemic, they were the only form of PHC in the country [4].

The first part of this paper analyzes the literature focusing on TAM in relation to the research hypotheses. The second part contains a description of the research methodology used in this study, with emphasis on the validation of the research tool developed for measuring the modified elements of TAM. The following section presents the results regarding the analyzed constructs of the model and the results of structural equation modeling (SEM). The last part of the paper contains the discussion, conclusions and practical implications.

Literature Review

Until March 2020, the use of telemedicine was negligible and mainly concerned patients in rural areas. In 2019, teleconsultations accounted for as much as 8% of all medical visits in the United States [5]. The situation has changed significantly after the outbreak of the COVID-19 pandemic, especially in the case of emergency visits. In the United States, there was a 683% increase in teleconsultations between March 2 and April 14, 2020 [5]. Researchers agree that the COVID-19 pandemic has changed the way we think about telemedicine. First and foremost, it popularized this way of providing medical services. However, its primary advantage is to protect both patients and physicians from the risk of virus infection [6,7]. In the long term, researchers believe that telemedicine can successfully replace a significant portion of face-to-face (F2F) visits; however, it will not eliminate them completely [4,8,9].

Telemedicine is not a new concept. Decades ago, pioneering projects emerged to test the concept of telemedicine or evaluate its applicability. However, most of these telemedicine projects failed to meet expectations. The failure was blamed on an underdeveloped and mostly primitive IT infrastructure, immature technology, and ineffective use [1]. The failure of first-generation telemedicine projects prompted an in-depth analysis and rigorous assessment of the technological, social, cultural, and organizational dimensions surrounding their introduction.

Resistance from users of new technologies in the medical community is as natural as possible, even if users are aware of the benefits this technology brings with it [10]. Therefore, an essential organizational challenge facing health care organizations considering or planning to provide telemedicine-enabled health care services is technology acceptance by users [1]. The problem regarding acceptance of technology has been discussed for a long time [11], and many models have been developed to assess users' attitudes toward new solutions [12-15]. Acceptance of ICT by physicians providing health care has also been assessed [16-18].

TAM is the most popular model in the literature for testing the acceptance of technology [19]. Van Schaik et al [20] used TAM to assess the attitudes of physiotherapists toward new medical technologies. Chau and Hu [21] studied the acceptance of telemedicine technologies by physicians. Holden and Karsh [22] extensively reviewed the literature on TAM applications and related models for ICT acceptance by health care professionals and, more specifically, health care information systems. They noted that in health care, there is a need for a complete approach to technology acceptance testing than for other professionals from companies or ICT organizations [22]. We chose to use TAM in our study because it is general, parsimonious, and ICT specific. It is designed to provide an explanation and prediction of the acceptance of a wide range of ICTs by a diverse population of users in different organizational contexts. In addition, the model has a well-researched and validated list of psychometric measurements, and this makes its use operationally attractive. Finally, TAM is the dominant model for studying user acceptance of technology, and over the years, it has accumulated satisfactory empirical support for its overall explanatory power and assumed individual causal relationships [1,2]. Over the past few decades, many researchers have proven that TAM enriched with certain other constructs is better suited to research and explain the acceptance of new information technologies by users [23].

TAM analyzes the influence of various factors on the intention to use new technology, among which the main role is played

XSL•FO
RenderX

by PU and PEU [24]. PU is defined as the extent to which the user's work performance is expected to improve through the use of new technology [2,25]. Similarly, Davis [11] defined PU as an individual's perceptions regarding the outcome of the experience with technology. In the area of health care, Kissi et al [26] defined PU as "physicians' belief regarding the benefits of telemedicine services that they improve access to medical care, the flow of medical records and patients' health."

PEU, in contrast, is the degree to which using new technology is expected to be effortless [2,25,27]. In the area of health care, PEU describes how physicians perceive telemedicine services in terms of their ease of use and learning [26].

TAM suggests that actual technology usage is determined by individuals' INT [2]. INT is understood as a motivation encouraging the system's user to use the system continuously, and in the case of physicians, it concerns their motivation to use telemedicine services, including, above all, teleconsultations [26]. INT is affected by PU, PEU, and users' attitudes toward technology [2,19]. Lin et al [28] used an integrated approach with the key elements from TAM and assessed the technology acceptance by health professionals of what they called "personal digital assistance (PDA)." The main variables from TAM (PU and self-efficacy) determined INT [28]. Similar conclusions were reached by Zayyad and Toycan [29], Chau and Hu [30], Tubaishat [31], and Vitari and Ologeanu-Taddei [32]. Thus, our first research hypotheses were proposed as follows:

- Hypothesis (H)1: PU has a direct effect on INT.
- H2: PEU has a direct effect on INT.

The results of Lin et al [28] showed that the traditional variables of TAM can be effectively integrated with variables from other theoretical approaches, which may help better understand the acceptance of new technologies by health care professionals [33]. There are relatively fewer TAM tests and modifications in the health care sector. Therefore, it is worth making such attempts to broaden the knowledge about new factors affecting physicians' acceptance of technologies. In our research, we decided to include 2 new constructs: SAT and Q. SAT explains how satisfied users are with using a particular service [19]; Q explains how physicians assess the value and worth of their work with the use of teleconsultations. The addition of these constructs was a consequence of both literature reviews and interviews with physicians (pilot survey). Bhattacherjee and Premkumar [10,34-36] noted that in the use phase of technology (postacceptance), PU is positively associated with user SAT. Alsohime et al [37] also confirmed the effect of PU on SAT, noting the significant impact of training courses before the implementation of new technology. Petter and Fruhling [38] confirmed that the SAT that users have with the information system positively affects their INT. As a consequence, we presented the subsequent hypotheses:

- H3: PU has a direct effect on SAT.
- H4: INT has a direct effect on SAT.

In our previous studies examining technology acceptance in Polish PHC facilities, we developed a conceptual framework defining the impact of PU and PEU on the need for teleconsultation adoption and examined the influence of selected behavioral factors on these constructs [2]. In this paper, we enriched these previous studies by analyzing the impact of PU and PEU as independent variables on SAT and Q. In our analysis, we additionally considered INT as a mediating variable in the model. This is the first study that extends TAM to include an analysis of SAT and Q in the teleconsultation condition.

Padilha et al [39] surveyed students and nurses' ease, usefulness, and intention to use a Massive Open Online Course. Findings confirmed the significant impact of PU, PEU, and INT on the current and future Q the groups studied [39]. Similar conclusions were reached by Saputra et al [40] and Chirchir et al [41]. Souza et al [42] proposed a process model for the evaluation of the Q of clinical decision support systems following the ISO/IEC 25022 and ISO/IEC 25010 standards, part of which was to identify the effect of SAT on Q. Given these considerations, we proposed the following hypotheses:

- H5: PU has a direct effect on Q.
- H6: PEU has a direct effect on Q.
- H7: INT has a direct effect on Q.
- H8: SAT has a direct effect on Q.

Although according to the research conducted so far, TAM is a reliable model for examining technology acceptance in PHC facilities, we can always try to supplement it with new research constructs [2]. TAM, in our study, was supplemented with 2 constructs, SAT and Q, which will contribute to the health care literature.

The Polish health care system is based on an insurance model. PHC physicians in Poland must be health insurance physicians who have a contract with the National Health Fund (NHF) to provide health care services. The functioning of PHCs in Poland is based on the right of patients to personally choose a preferred physician. The selected physicians receive an annual capitation fee for each registered patient [43]. PHC facilities in Poland function as both state-owned and private facilities. Both sign contracts to provide services that are free to the patient and paid for by the NHF. Each facility is managed according to its own rules. Private facilities have more flexibility in making decisions and in hiring and paying employees. Public facilities are subject to top-down regulations governing their operations. As private facilities provide fee-based services, in addition to free services, they have more resources to pay salaries and run their operations.

Methods

Data Collection

In this cross-sectional study, the survey followed a multimodal approach, integrating computer-assisted web interviewing (CAWI), computer-assisted telephone interviewing (CATI), and paper-and-pencil interviewing (PAPI) techniques across a statistically representative sample of 371 PHC facilities. This number was derived from a total of 5503 outpatient PHC facilities in Poland, calculated to be representative at a 95% CI level, with a 50% response distribution and a 5% margin of error, for the aforementioned assumptions, and the minimum survey sample size was 359 [44-47]. The survey sample was randomly selected from the BISNODE database, which includes

XSL•FO
RenderX

comprehensive information on all Polish PHC facilities. Of 5503 outpatient facilities in Poland, 371 (6.7%) were successfully surveyed, with each representing 1 physician providing remote medical advice. The survey process entailed replacing nonparticipating facilities with other randomly chosen facilities, ensuring the integrity and representativeness of the sample. Quality control measures were rigorously followed, with a certified polling company overseeing the survey execution. Instances of schematic responses and unusually short survey durations led to the exclusion and replacement of certain responses, resulting in a final analytical sample of 361 (97.3%) records. Before filling the questionnaire, the physicians were informed that the questionnaire is aimed at PHC physicians and concerns the evaluation of their satisfaction with the use of the teleconsultation system for the provision of patient care.

The sample was limited to 1 PHC physician from each randomly selected facility in Poland. This approach was adopted for several reasons. Conducting a survey that included multiple physicians from each facility would have significantly increased the scale and complexity of the study. Given resource constraints, such as funding, time, and personnel, it was more feasible to limit the number of participants, while still achieving a representative sample. The aim was to obtain a broad overview of the acceptance and satisfaction with teleconsultation across a wide range of PHC facilities in Poland. By selecting 1 physician from each facility, we ensured a diverse and statistically representative sample of the entire population of PHC facilities, which may not have been possible with a more concentrated sample from fewer facilities. The study was primarily designed to assess the impact of system-level factors (eg, PU and PUE) on the acceptance of telemedicine. Although individual characteristics, such as age and gender, are important, the primary focus was on broader systemic issues that could be generalized across the population. Conducting an extensive survey during the COVID-19 pandemic posed unique challenges, including limited access to PHC facilities and the need to minimize contact. The study design did not include a detailed examination of individual physician factors and their impact on the acceptance of telemedicine, and the approach was strategically chosen to balance comprehensiveness, feasibility, and the overarching research objectives.

Survey Instrument and Measures

The survey instrument contained 2 groups of statements and questions: statements about analyzed latent factors and general questions about age and gender of the respondent, legal status of the PHC facility, voluntariness of providing remote advice, and ways in which the respondent provided remote advice. The questionnaire included 48 statements and questions, but only the statements used in the modified TAM are presented in Multimedia Appendix 1. A 5-point Likert scale was used in this study: 1 (I do not agree), 2 (I do not agree somewhat), 3 (I neither agree nor disagree), 4 (I agree somewhat), and 5 (I agree).

SEM in this study was adapted from the original TAM, which identifies PU and PEU as the principal determinants of technology use and acceptance. The PU variable measures the degree to which physicians believe that teleconsultations improve their work efficiency and patient care. It encompasses aspects such as enhanced health care delivery, better documentation, and cost-effective monitoring. PU in this context is gauged through 6 survey statements (PU1-PU6), which assess various dimensions of utility that teleconsultations provide in the health care setting [11,48-51]. PEU variables are defined as the extent to which physicians believe that teleconsultations are effortless to learn and implement. The variables relate to the ease of use of the teleconsultation system and the use of medical data. This construct is evaluated via 7 survey statements (PEU1-PEU7), focusing on the usability and accessibility of the teleconsultation system [2,25,27]. INT, similar to that in TAM, used as a mediating variable, represents the likelihood of physicians continuing to use teleconsultations in the future. It is measured through 5 survey statements (INT1-INT5), focusing on the perceived long-term utility and effectiveness of teleconsultations in patient diagnosis and care [11,12,41]. Q, used as an independent variable, refers to the perceived enhancement in work value and worth due to teleconsultations and is gauged through 3 survey statements (Q1-Q3). It assesses whether teleconsultations uphold the standard of traditional visits and enable comprehensive patient care [22]. The SAT variable measures the overall contentment of physicians with teleconsultations. It includes aspects such as convenience compared to traditional visits and comfort in providing remote advice [11,52] and is assessed through 4 survey statements (SAT1-SAT4) [12,13,53,54].

Exploratory Factor Analysis

The study required validating the structure and dynamics within the adapted TAM, because the original model was extended to encompass SAT and Q. The evaluation of the survey statements for inclusion in the factors measuring the assessed dimensions was based on exploratory factor analysis (EFA). EFA was used to select the final variables for the structural model. For each dimension, EFA was separately carried out to assess the Kaiser-Meyer-Olkin (KMO) measure of sampling adequacy. The value of this index should be >0.7 [55]. For all dimensions, this condition was fulfilled, but the KMO value of the PEU dimension was <0.7. The EFA results for each dimension are presented in Table 1.

In addition, the ability of each dimension to be represented by individual survey statements was assessed using Bartlett's test of sphericity. For each dimension, the chi-square value was significant, and in each case, $P<.001$. Based on EFA, the PEU dimension was finally divided into 2 separate factors: PEU_1 and PEU_2. Questions PEU1, PEU2, PEU3, and PEU4 were about the technical ease of use of the system (PEU_1), and questions PEU5, PEU6, and PEU7 were about the ease of use of the system from the point of view of handling medical data (PEU_2). Reliability analysis was conducted for all dimensions of the validity of using the adopted statements to measure each factor. Cronbach coefficients were determined for each factor, with acceptable values falling within the range of 0.7-0.95 [56]. The constructs were confirmed to possess suitable psychometric properties, enabling their effective use in SEM analysis. For the PEU dimension, reliability analysis did not give a clear answer as to which statement should be removed to improve the Cronbach and KMO coefficient values. The use of survey

XSL•FO
RenderX

statements to measure the PEU dimension requires confirmation in the structural model. The EFA results for the PEU dimension are presented in Table 2, and component factor loadings are presented in Table 3.

Based on imputed factors, a structural model was prepared, where the effects of PU, PEU_1, and PEU_2 on INT, SAT, and Q were studied. Figure 1 shows the final tested model with only significant dependencies.

Table 1. Component factor loadings.[a]

Variable	Factor					
	PU[b]	PEU[c]_1	PEU_2	INT[d]	SAT[e]	Q[f]
KMO[g]	0.90	0.69	0.69	0.80	0.77	0.73
Cronbach	0.89	0.71	0.71	0.81	0..89	0.86
PU1	0.69	__[h]	—	—	—	—
PU2	0.84	—	—	—	—	—
PU3	0.89	—	—	—	—	—
PU4	0.87	—	—	—	—	—
PU5	0.79	—	—	—	—	—
PU6	0.79	—	—	—	—	—
PEU1	—	0.79	—	—	—	—
PEU2	—	0.69	—	—	—	—
PEU3	—	0.59	—	—	—	—
PEU4	—	0.81	—	—	—	—
PEU5	—	—	0.67	—	—	—
PEU6	—	—	0.86	—	—	—
PEU7	—	—	0.78	—	—	—
INT1	—	—	—	0.80	—	—
INT3	—	—	—	0.82	—	—
INT4	—	—	—	0.76	—	—
INT5	—	—	—	0.83	—	—
SAT1	—	—	—	—	0.78	—
SAT2	—	—	—	—	0.88	—
SAT3	—	—	—	—	0.89	—
SAT4	—	—	—	—	0.82	—
Q1	—	—	—	—	—	0.86
Q2	—	—	—	—	—	0.90
Q3	—	—	—	—	—	0.90

[a]Extraction method: principal component analysis; rotation method: varimax with Kaiser normalization.
[b]PU: perceived usefulness.
[c]PEU: perceived ease of use.
[d]INT: intention to use teleconsultation.
[e]SAT: satisfaction.
[f]Q: quality of work.
[g]KMO: Kaiser-Meyer-Olkin.
[h]Not applicable.

XSL•FO
RenderX

Table 2. Total variance of the PEU[a] dimension explained by EFA[b,c].

Component	Initial eigenvalues			Extraction sums of squared loadings			Rotation sums of squared loadings		
	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %
1	2.71	38.68	38.68	2.71	38.68	38.68	2.20	31.40	31.40
2	1.40	19.98	58.65	1.40	19.98	58.65	1.91	27.25	58.65
3	0.90	12.86	71.52	__[d]	—	—	—	—	—
4	0.69	9.83	81.35	—	—	—	—	—	—
5	0.58	8.34	89.69	—	—	—	—	—	—
6	0.37	5.28	94.96	—	—	—	—	—	—
7	0.35	5.04	100.00	—	—	—	—	—	—

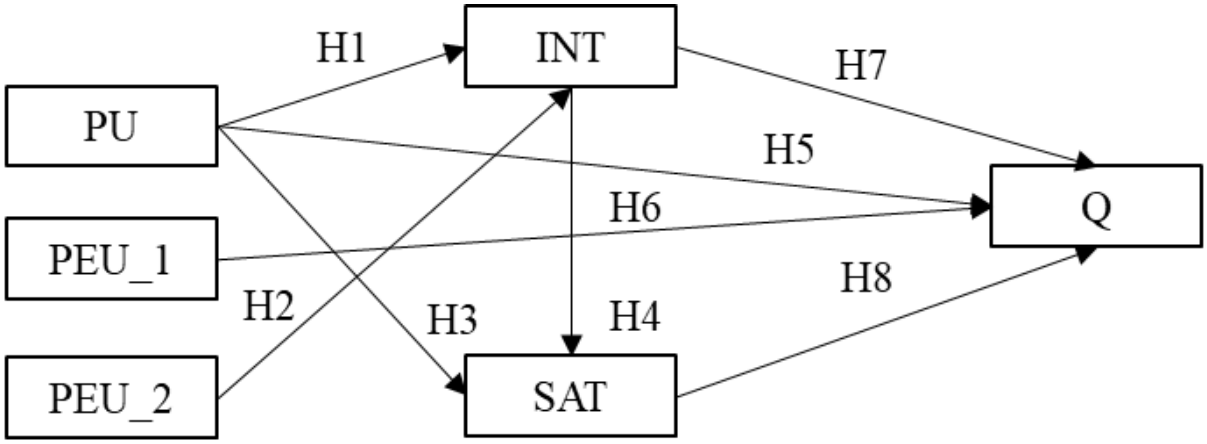
[a]PEU: perceived ease of use.
[b]EFA: exploratory factor analysis.
[c]Extraction method: principal component analysis.
[d]Not applicable.

Table 3. PEU[a] factors loadings.[b]

Variable	Factor	
	PEU_1	PEU_2
PEU1	0.787	__[c]
PEU2	0.685	—
PEU3	0.589	—
PEU4	0.808	—
PEU5	—	0.670
PEU6	—	0.863
PEU7	—	0.784

[a]PEU: perceived ease of use.
[b]Extraction method: principal component analysis; rotation method: varimax with Kaiser normalization.
[c]Not applicable.

Figure 1. Structural model depicting the factors affecting SAT with teleconsultation and Q. Research hypotheses H1-H8 for direct paths. H: hypothesis; INT: intention to use teleconsultation; PEU: perceived ease of use; PU: perceived usefulness; Q: quality of work; SAT: satisfaction.



In the original TAM, the actual use of technology is the independent model variable. In this study, physicians were required to conduct teleconsultations, so INT and SAT variables were treated as mediators affecting the dependent variable, Q. Because of this, the effect of PU, PEU1, and PEU_2 variables on Q was also studied indirectly. The research hypotheses H9-H14 regarding indirect effects were proposed as follows:

- H9: PU has an indirect effect on Q through INT.
- H10: PU has an indirect effect on SAT through INT.

XSL•FO
RenderX

- H11: PU has an indirect effect on Q through INT and SAT.
- H12: PEU_2 has an indirect effect on Q through INT.
- H13: PEU_2 has an indirect effect on SAT through INT.
- H14: PEU_2 has an indirect effect on Q through INT and SAT.

Ethical Considerations

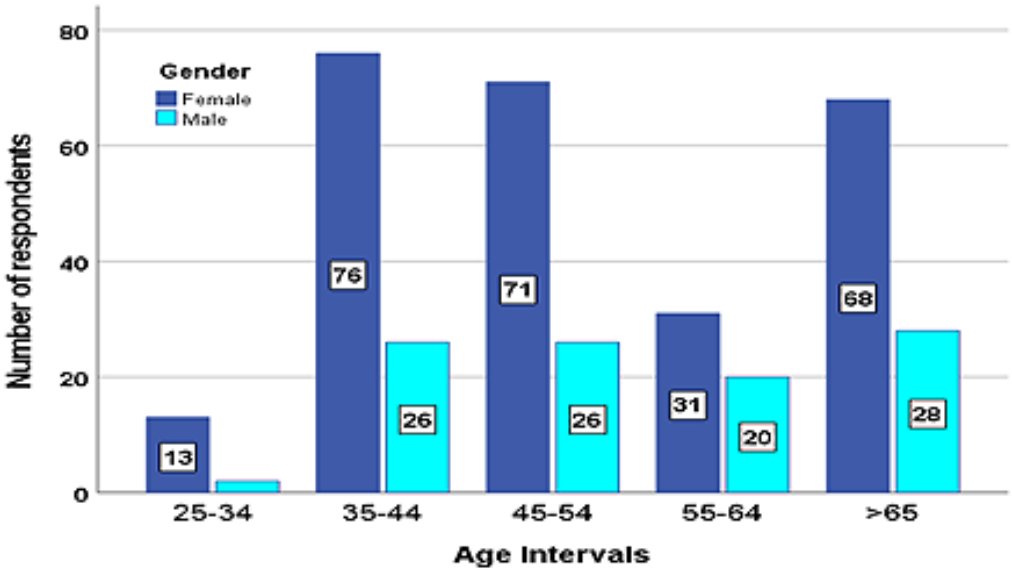
The survey instrument was approved by the Ethics Committee of the Warsaw University of Technology that issued the Certificate of Ethics Approval (certificate dated January 15, 2021). As a result of the contact, 587 physicians provided consent to participate in the study, of which, despite consent, in 216 (36.8%) PHC facilities, the complete set of surveys could not be completed. Respondents were informed about the purpose of the research before starting the survey. They could withdraw from completing the survey at any time. Study data were anonymous and deidentified. No compensation from respondents was taken for the research.

Results

Participant Characteristics

Of the 361 physicians, 199 (55.1%) were in the 35-54–year age group, 94 (26%) were of retirement age and over 65 years old, 260 (72%) were women, and only 101 (28%) were men. The age distribution of the surveyed physicians is presented in Figure 2.

Figure 2. Physicians' age distribution by gender.



Evaluation of the Level of TAM Dimensions

The research considered the dimensions originally defined in TAM (PU, PEU, and INT) [11]. Two dimensions were added to the model: SAT and Q.

The PU variable assesses physicians' perceptions of the utility and benefits of teleconsultation services. The mean scores for PU1-PU6 ranged from 3.85 to 4.36 (SD 0.87-1.18), indicating a generally high level of agreement among physicians that teleconsultations are beneficial to their work. The highest mean score was for PU1 (4.36, SD 0.94), suggesting that the physicians particularly valued teleconsultations during challenging times, such as pandemics. The SDs, ranging from 0.87 (mean 4.24) for PU4 to 1.18 (mean 3.85) for PU3, implied some variability in how the physicians perceived the usefulness of different aspects of teleconsultations. The higher deviation in PU3 indicated more varied opinions about the efficiency enhancement brought by teleconsultations. The skewness for all PU statements was negative, ranging from –0.81 to –1.73, suggesting a tendency among the physicians to agree that teleconsultations are useful in their work. The most pronounced skewness was observed in PU1, indicating strong agreement about the difficulty of work during a pandemic without teleconsultations. The kurtosis values ranged from –0.38 for PU5 to 2.77 for PU1, indicating varied distribution patterns. The higher kurtosis in PU1 reflected a more peaked distribution, suggesting more consistent agreement among physicians regarding its statement. The data suggested that physicians perceive teleconsultations as a valuable tool in their professional practice. The high mean scores across all PU items reflected positive perceptions of the teleconsultation system's usefulness, particularly in aiding work during a pandemic (PU1). The variation in SDs pointed to some differences in individual opinions about the specific benefits of teleconsultations, such as work efficiency and time-saving aspects. The negative skewness across all items highlighted a general agreement on the usefulness of teleconsultations, with a stronger consensus in areas such as coping with pandemic challenges. The kurtosis values, particularly for PU1, indicated that most responses were concentrated around higher ratings, showing strong agreement in specific areas of usefulness. The PU variable demonstrated that physicians generally regard teleconsultations as a beneficial tool in their practice, particularly under challenging conditions, such as a pandemic. Although there was overall agreement on their utility, the variation in perceptions across different aspects suggests areas where experiences and expectations of teleconsultations may differ among individual physicians.

The PEU variable examines physicians' perceptions of the ease and effortlessness associated with using teleconsultation systems. The mean scores for PEU1-PEU7 ranged from 3.81

XSL•FO
RenderX

to 4.60 (SD 0.60-1.42), respectively, and indicated a generally high level of agreement among the physicians that teleconsultation systems are user friendly and easy to use. The highest mean score was for PEU7 (4.60, SD 0.60), suggesting that using external systems during teleconsultations was perceived as particularly straightforward by most of the physicians. The SDs, ranging from 0.60 (mean 4.60) for PEU7 to 1.42 (mean 3.81) for PEU2, suggested a variation in the perceptions of ease of use, with the greatest variation in responses relating to the intellectual effort required (PEU2). The skewness for all PEU statements was negative, ranging from –0.99 to –1.99. This indicated a tendency among the physicians to rate the ease of use of teleconsultation systems highly, with a pronounced leaning toward agreement for most statements, particularly PEU7. The kurtosis values for the PEU variables were varied, with some (eg, PEU7) indicating a highly peaked distribution (7.13), which suggests that responses were more consistently clustered around the higher end of the scale. Overall, the data indicated that physicians find teleconsultation systems relatively easy to use. The high mean scores across all PEU items reflected a positive perception of the teleconsultation system's usability. The variation in SDs, especially the higher deviation for PEU2, suggested that although using the teleconsultation system is generally perceived as easy to use, there are aspects, such as the intellectual effort required, where opinions vary more widely. The pronounced negative skewness, especially for items such as PEU7, underlined a strong agreement in the ease of integrating and using external systems, which might be due to prior familiarity and necessity in clinical practice. The high kurtosis value for PEU7 pointed to a strong consensus among the respondents about the ease of this particular aspect of teleconsultation systems. The physicians generally found teleconsultation systems to be user friendly, with some areas, such as integration with external systems, being particularly well received. Around 253-325 (70.1%-90%) respondents rated the questions in the PEU group positively. However, there was notable variability in perceptions regarding the intellectual effort required, suggesting areas for potential improvement or further training.

The INT variable reflects physicians' intentions and willingness to continue using teleconsultation services in the future. The mean scores for INT1, INT3, INT4, and INT5 were 4.22 (SD 0.89), 3.99 (SD 1.01), 3.87 (SD 1.12), and 4.06 (SD 0.94), respectively. These scores, being close to or above 4 on a 5-point scale, suggested a generally positive inclination among the physicians toward the continued use of teleconsultations. The SDs for INT1, INT3, INT4, and INT5 indicated a moderate level of variation in responses. This variation signified that although there is a general trend of positive intention, individual physicians' perspectives on the future use of teleconsultations vary. The skewness values for these variables ranged from –0.97 to –1.20, which are negative. This negative skewness indicated a tendency among respondents to agree with the statements related to INT, suggesting that a larger proportion of physicians are inclined to continue using these services. The kurtosis values for INT1 (1.26), INT3 (0.46), INT4 (0.22), and INT5 (0.79) showed a relatively normal to slightly peaked distribution. This indicated a consistent pattern in physicians' responses, with a tendency toward agreement on the future use of teleconsultations. The data indicated a positive attitude among physicians toward continuing the use of teleconsultations. The inclination to add video consultations to telephone conversations and to use teleconsultations for patient diagnosis and collaboration with other physicians was evident. The moderate spread in responses, however, pointed to some differences in enthusiasm or confidence about teleconsultations among individual physicians. These differences could be attributed to factors such as personal experience, familiarity with technology, or specific demands of their medical practice. Around 253-289 (70.1%-80%) of respondents intend to use teleconsulting in the future.

The Q variable was rated by physicians as the lowest level of all dimensions. The mean scores ranged from 3.28 (SD 1.26) to 3.73 (SD 1.06). The spread of responses was not even. The mean scores for Q1, Q2, and Q3 were 3.73 (SD 1.06), 3.43 (SD 1.19), and 3.28 (SD 1.26), respectively. These scores indicated a moderate level of agreement among physicians regarding Q, with some variability in the perception of different aspects of Q. The SDs for Q1-Q3 suggested a significant spread in responses. This indicated a varied perception of Q among the physicians, reflecting diverse experiences and expectations regarding teleconsultations. The skewness values for Q1 (–0.82), Q2 (–0.55), and Q3 (–0.27) were negative, implying a tendency for responses to lean toward agreeing with the statements about Q, although this tendency was less pronounced compared to other variables, such as SAT or PU. The kurtosis values for Q1 (–0.01), Q2 (–0.82), and Q3 (–1.18) suggested a relatively flat distribution, particularly for Q2 and Q3. This flatness indicated that the responses were more evenly spread across the scale, reflecting a wide range of opinions on the quality of work. The data suggested that although there is a general trend of moderate satisfaction with Q, there is considerable variation in how physicians perceive this Q. This variability could be influenced by different factors, such as the type of teleconsultation services used, the technological infrastructure in place, and the specific needs of the patient population being served. The more even distribution of responses, especially for Q2 and Q3, indicated that opinions on Q are diverse. This diversity might reflect the complexity of evaluating health care quality in a remote setting, where factors such as patient interaction, diagnostic accuracy, and treatment effectiveness play a crucial role. Only 199 (55.1%) respondents felt that teleconsultations are suitable for holistic patient care, 130 (36%) felt that remote consultations do not allow for comprehensive care, 108 (29.9%) felt that the Q by both methods is not the same, and 224 (62%) believed that the Q using remote visits is similar to that of visits in the clinic. Most physicians (n=260, 72%) agreed with the statement that teleconsultations improve Q.

The SAT variable offers valuable insights into physicians' contentment and approval levels regarding the use of teleconsultations. Physicians positively rated statements regarding SAT. The mean scores for the 4 statements of the SAT variable (SAT1, SAT2, SAT3, and SAT4) were 3.55 (SD 1.16), 4.13 (SD 0.88), 4.00 (SD 0.99), and 3.91 (SD 1.01), respectively. These scores, hovering around or above 4 on a 5-point scale, indicated a general trend of SAT among physicians with teleconsultation services. The SDs for

XSL•FO
RenderX

SAT1-SAT4 suggested a moderate level of variation in responses. This indicated that although there is an overall sense of SAT, there are differences in individual experiences and perceptions regarding teleconsultation services. The skewness values for these variables ranged from –0.62 to –1.14, which are negative. This negative skewness implied a leaning toward higher SAT ratings among the respondents, indicating that more physicians agree with the positive aspects of teleconsultations. The kurtosis values for SAT1 (–0.52), SAT2 (1.51), SAT3 (0.82), and SAT4 (0.78) suggested a mixed distribution pattern. Although SAT1 indicated a relatively normal distribution, SAT2 showed a more peaked distribution, suggesting more consistent high ratings among physicians. The data suggested that physicians are generally satisfied with teleconsultation services, as indicated by the mean scores leaning toward the higher end of the scale. The moderate variation in responses indicated differing levels of SAT, which may be influenced by individual experiences, technological proficiency, or specific needs in their practice. The skewness toward higher SAT ratings suggested that a larger proportion of physicians find teleconsultations to be a convenient and effective medium for providing health care services.

Descriptive statistics of all model variables are presented in Table 4, and distributions of the responses are presented in Table 5.

Table 4. Descriptive statistics of model variables.

	Mean (SD)	Variance	Skewness	Kurtosis
PU1[a]	4.36 (0.94)	0.87	–1.73	2.77
PU2	4.04 (1.03)	1.06	–1.25	1.07
PU3	3.85 (1.18)	1.38	–0.91	–0.14
PU4	4.24 (0.87)	0.75	–1.34	1.93
PU5	3.89 (1.14)	1.30	–0.81	–0.38
PU6	4.03 (1.01)	1.01	–1.16	0.99
PEU1[b]	4.28 (0.83)	0.70	–1.34	1.88
PEU2	3.81 (1.42)	2.02	–0.99	–0.47
PEU3	4.51 (0.67)	0.45	–1.82	5.27
PEU4	4.06 (1.04)	1.08	–1.27	1.14
PEU5	4.19 (0.88)	0.77	–1.63	3.44
PEU6	4.42 (0.69)	0.48	–1.48	3.82
PEU7	4.60 (0.60)	0.36	–1.99	7.13
INT1[c]	4.22 (0.89)	0.80	–1.20	1.26
INT3	3.99 (1.01)	1.01	–0.97	0.46
INT4	3.87 (1.12)	1.25	–0.97	0.22
INT5	4.06 (0.94)	0.88	–1.04	0.79
Q1[d]	3.73 (1.06)	1.13	–0.82	–0.01
Q2	3.43 (1.19)	1.42	–0.55	–0.82
Q3	3.28 (1.26)	1.58	–0.27	–1.18
SAT1[e]	3.55 (1.16)	1.35	–0.62	–0.52
SAT2	4.13 (0.88)	0.78	–1.14	1.51
SAT3	4.00 (0.99)	0.98	–1.02	0.82
SAT4	3.91 (1.01)	1.02	–1.07	0.78

[a]PU: perceived usefulness.

[b]PEU: perceived ease of use.

[c]INT: intention to use teleconsultation.

[d]Q: quality of work.

[e]SAT: satisfaction.

XSL•FO
RenderX

Table 5. Participant (N=361) response distributions.

Likert scale response	I do not agree, n (%)	I do not agree somewhat, n (%)	I neither agree nor disagree, n (%)	I agree somewhat, n (%)	I agree, n (%)
PU1[a]	7 (1.9)	19 (5.3)	17 (4.7)	113 (31.3)	205 (56.8)
PU2	11 (3.0)	32 (8.9)	22 (6.1)	162 (44.9)	134 (37.1)
PU3	18 (5.0)	45 (12.5)	37 (10.2)	134 (37.1)	127 (35.2)
PU4	4 (1.1)	17 (4.7)	27 (7.5)	152 (42.1)	161 (44.6)
PU5	10 (2.8)	51 (14.1)	43 (11.9)	123 (34.1)	134 (37.1)
PU6	10 (2.8)	27 (7.5)	36 (10.0)	157 (43.5)	131 (36.3)
PEU1[b]	2 (0.6)	19 (5.3)	20 (5.5)	156 (43.2)	164 (45.4)
PEU2	48 (13.3)	35 (9.7)	11 (3.0)	111 (30.7)	156 (43.2)
PEU3	2 (0.6)	6 (1.7)	6 (1.7)	138 (38.2)	209 (57.9)
PEU4	12 (3.3)	29 (8.0)	25 (6.9)	154 (42.7)	141 (39.1)
PEU5	9 (2.5)	14 (3.9)	14 (3.9)	186 (51.5)	138 (38.2)
PEU6	2 (0.6)	6 (1.7)	12 (3.3)	160 (44.3)	181 (50.1)
PEU7	2 (0.6)	2 (0.6)	4 (1.1)	123 (34.1)	230 (63.7)
INT1[c]	4 (1.1)	16 (4.4)	41 (11.4)	136 (37.7)	164 (45.4)
INT3	8 (2.2)	28 (7.8)	52 (14.4)	145 (40.2)	128 (35.5)
INT4	17 (4.7)	35 (9.7)	45 (12.5)	146 (40.4)	118 (32.7)
INT5	5 (1.4)	25 (6.9)	43 (11.9)	157 (43.5)	131 (36.3)
Q1[d]	13 (3.6)	47 (13.0)	47 (13.0)	172 (47.6)	82 (22.7)
Q2	25 (6.9)	78 (21.6)	33 (9.1)	167 (46.3)	58 (16.1)
Q3	29 (8.0)	99 (27.4)	36 (10.0)	136 (37.7)	61 (16.9)
SAT1[e]	23 (6.4)	55 (15.2)	60 (16.6)	148 (41.0)	75 (20.8)
SAT2	6 (1.7)	12 (3.3)	48 (13.3)	157 (43.5)	138 (38.2)
SAT3	10 (2.8)	19 (5.3)	59 (16.3)	146 (40.4)	127 (35.2)
SAT4	11 (3.0)	33 (9.1)	36 (10.0)	178 (49.3)	103 (28.5)

[a]PU: perceived usefulness.

[b]PEU: perceived ease of use.

[c]INT: intention to use teleconsultation.

[d]Q: quality of work.

[e]SAT: satisfaction.

Structural Equation Modeling

Based on the factors extracted in EFA, a structural model was developed (Figure 3). The model demonstrated excellent fit with the data, as indicated by indices such as the chi-square-to-degrees-of-freedom index: PCMIN/DF=0.91 (<5 is acceptable), comparative fit index (CFI)=1 (>0.9 is acceptable), goodness-of-fit index (GFI)=0.99 (>0.9 is acceptable), and root mean square error of approximation (RMSEA)=0 (<0.08 is acceptable). These values suggested that our model is robust and accurately represents the observed data.

In assessing the relationships between key constructs in our teleconsultation acceptance model, we used maximum likelihood estimates to derive regression weights (Table 6). Our analysis revealed significant relationships between the constructs, as evidenced by the *P* values and critical ratios (CRs) in the regression weights. The influence of PU on INT was strong (estimate=0.63, CR=15.84, *P*<.001), suggesting that physicians' PU of teleconsultations significantly predicts their INT. PEU_2 also positively influenced INT (estimate=0.17, CR=4.31, *P*<.001), albeit to a lesser extent than PU. Both PU (estimate=0.44, CR=9.53, *P*<.001) and INT (estimate=0.4, CR=8.57, *P*<.001) significantly predicted SAT, indicating that PU and INT are crucial determinants of SAT with teleconsultations. Q was positively influenced by INT (estimate=0.179, CR=3.64, *P*<.001), PU (estimate=0.246, CR=4.79, *P*<.001), PEU_1 (estimate=0.18, CR=4.93, *P*<.001), and SAT (estimate=0.357, CR=6.97, *P*<.001), highlighting a multifaceted impact on Q. Since on each path of the model, the values of the regression parameters had *P*<.05, it can be said

XSL•FO
RenderX

that at a significance level of =.05, all the direct dependencies of the model are significant, and hence, H1-H8 are supported.

Standardized regression weights underscore the relative strength of these relationships in standardized format, which is particularly useful for comparing the effects across different predictors within our model. A significant positive covariance was observed between PU and PEU_1 (estimate=0.37, CR=6.52, *P*<.001) and between PU and PEU_2 (estimate=0.28, CR=5.05, *P*<.001), suggesting that PU and PEU are interrelated constructs. However, no significant correlation was found between PEU_1 and PEU_2 (estimate=0, *P*=.99), indicating these aspects of PEU may independently influence the model. The squared multiple correlations for INT (0.48), SAT (0.58), and Q (0.6) indicate a substantial proportion of variance in these endogenous variables, explained by their respective predictors.

Our findings provide strong empirical support for the hypothesized relationships within TAM. The significant paths between constructs such as PU, PEU, INT, SAT, and Q highlight the multifaceted nature of teleconsultation acceptance among physicians.

These results underscore the importance of both PU and PEU in influencing INT and SAT, which, in turn, impact Q.

Figure 3. Structural model presenting standardized estimates. BI: measurement error; INT: intention to use teleconsultation; PEU: perceived ease of use; PU: perceived usefulness; Q: quality of work; SAT: satisfaction.

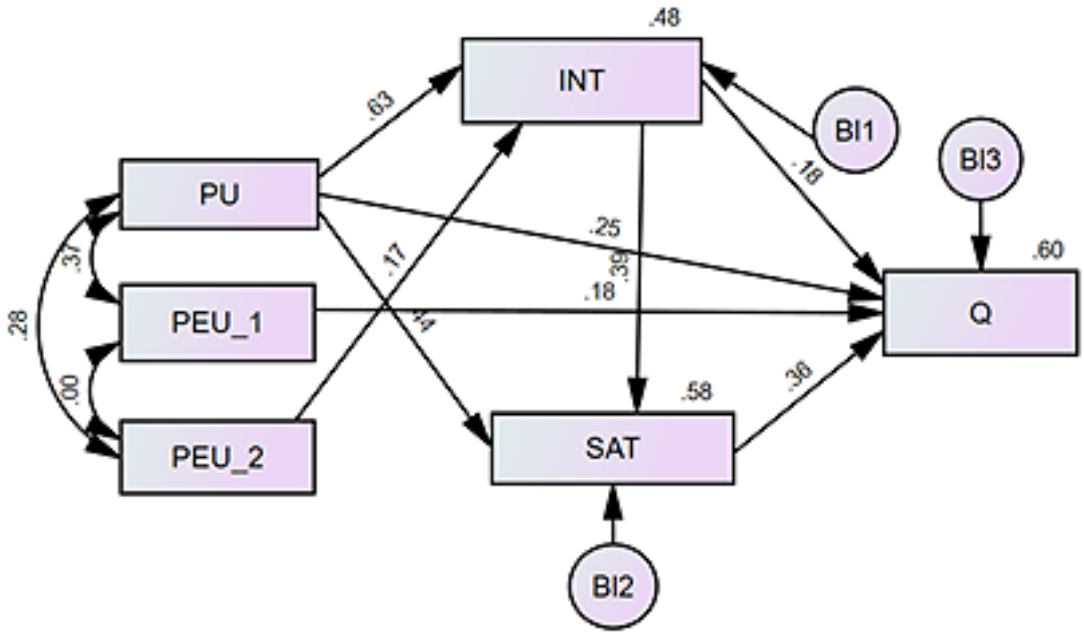


Table 6. Standardized regression weights of direct model paths.

Model paths	Estimate	*P* value
PU[a]→INT[b]	0.626	<.001
PEU[c]_2→INT	0.17	<.001
PU→SAT[d]	0.439	<.001
INT→SAT	0.395	<.001
INT→Q[e]	0.179	<.001
PU→Q	0.246	<.001
PEU_1→Q	0.176	<.001
SAT→Q	0.357	<.001

[a]PU: perceived usefulness.

[b]INT: intention to use teleconsultation.

[c]PEU: perceived ease of use.

[d]SAT: satisfaction.

[e]Q: quality of work.

XSL•FO
RenderX

Mediation Analysis

In addition to the analysis of the model's direct paths, a mediation analysis was conducted. The values of regression factors for indirect model paths were determined (Table 7). The mediating effects within the structural model for the teleconsultation research provide insights into the indirect pathways through which independent variables influence dependent variables. PU had a substantial indirect effect on SAT through INT, with an estimate of 0.25 (P<.001). This suggests that the PU of teleconsultation systems significantly influences SAT, mediated by INT. Additionally, PU indirectly affected Q through both INT (estimate=0.112, P<.002) and SAT (estimate=0.16, P<.001). These paths indicate that PU leads to higher Q, as it influences INT and SAT with it. PEU_2 indirectly influenced SAT via INT, with an estimate of 0.07 (P<.001). This effect signifies that the PEU contributes to SAT through INT. For Q, PEU_2 had an indirect effect through the mediation of INT (estimate=0.03, P<.001) and SAT (estimate=0.02, P<.001). These findings imply that PEU of teleconsultations not only impacts INT but also enhances the Q delivered through SAT. The path from INT to Q mediated by SAT was significant, with an estimate of 0.14 (P<.001), indicating that INT contributes to Q, as mediated by SAT levels. The bootstrap 95% CIs for these indirect effects reinforced their significance, as they did not include 0, and the P values were well below the .001 threshold, indicating robustness in these mediating relationships. The mediating effects elucidate the important role of INT and SAT with teleconsultation services in enhancing the PU and PEU.

Our findings from the SEM analysis corroborate several key hypotheses concerning direct, indirect, and mediating effects within the teleconsultation context. In accordance with H1, PU exhibited a significant direct effect on INT. Supporting H2, PU was also found to directly affect Q. In line with H3, a direct effect of PU on SAT was substantiated. PEU_1's direct effect on Q affirmed H4, demonstrating the technical influence of PEU on Q. PEU_2 was confirmed to directly impact INT, lending credence to H5. H6, positing a direct effect of INT on SAT, was also validated. Similarly, H7 was supported, with INT having a direct influence on Q. SAT was found to directly affect Q, confirming H8. The indirect influence of PU on Q through INT, posited in H9, was also substantiated. PU was found to indirectly affect SAT through INT, supporting H10. Moreover, the hypothesized indirect effect of PU on Q via the mediating roles of INT and SAT, as stated in H11, was validated. PEU_2's indirect impact on Q through INT, as hypothesized in H12, was confirmed. The indirect effect of PEU_2 on SAT via INT, detailed in H13, was likewise corroborated (Table 8).

Since on each path of the model, the values of the regression parameters had P<.05, it can be said that at a significance level of =.05, all the indirect dependencies of the model are significant, and hence, H9-H14 are supported.

Table 7. Standardized regression weights of indirect model paths.

Indirect model paths	Estimate	P value	Significance
PU[a]→INT[b]→SAT[c]	0.247	<.001	Significant
PU→INT→SAT→Q[d]	0.088	<.001	Significant
PU→INT→Q	0.112	<.002	Significant
PU→SAT→Q	0.157	<.001	Significant
PEU[e]_2→INT→SAT	0.067	<.001	Significant
PEU_2→INT→SAT→Q	0.024	<.001	Significant
PEU_2→INT→Q	0.03	<.001	Significant
INT→SAT→Q	0.141	<.001	Significant

[a]PU: perceived usefulness.

[b]INT: intention to use teleconsultation.

[c]SAT: satisfaction.

[d]Q: quality of work.

[e]PEU: perceived ease of use.

XSL•FO
RenderX

Table 8. Corroboration of hypotheses concerning direct, indirect, and mediating effects.

Hypothesis (H)	Description	Supported (yes/no)
H1	PU[a] has a direct effect on INT[b].	Yes
H2	PEU[c] has a direct effect on INT.	Yes
H3	PU has a direct effect on SAT[d].	Yes
H4	INT has a direct effect on SAT.	Yes
H5	PU has a direct effect on Q[e].	Yes
H6	PEU has a direct effect on Q.	Yes
H7	INT has a direct effect on Q.	Yes
H8	SAT has a direct effect on Q.	Yes
H9	PU has an indirect effect on Q through INT.	Yes
H10	PU has an indirect effect on SAT through INT.	Yes
H11	PU has an indirect effect on Q through INT and SAT.	Yes
H12	PEU_2 has an indirect effect on Q through INT.	Yes
H13	PEU_2 has an indirect effect on SAT through INT.	Yes
H14	PEU_2 has an indirect effect on Q through INT and SAT.	Yes

[a]PU: perceived usefulness.
[b]INT: intention to use teleconsultation.
[c]PEU: perceived ease of use.
[d]SAT: satisfaction.
[e]Q: quality of work.

Discussion

Principal Findings

This study evaluated Polish physicians' acceptance of teleconsultations during the COVID-19 pandemic in Poland. Most of the physicians positively assessed the PU of teleconsultations. The majority of physicians believed that their work during the COVID-19 pandemic would have been difficult without teleconsultations (88%) and that teleconsultations turned out to be a useful system enabling medical care (87%). The least number of physicians said that teleconsultations save time (61%) and improve performance (72%). Physicians are willing to use new technologies if they do not require additional time and effort, which is in line with other studies [25]. Similar results regarding the usefulness of teleconsultations during a pandemic and the ease of using them were obtained in a cross-sectional study conducted in 2020 in one of the Romanian counties using a questionnaire that assessed, among other things, the perception of teleconsultations by physicians. The study showed a positive perception of telemedicine by Romanian physicians. However, the researchers also highlighted the cons of teleconsultations, such as the time-consuming process, fear of making medical errors remotely, and communication difficulties on the part of patients [57]. The time-consuming nature of teleconsultations has also been confirmed in Great Britain; British physicians reported on time-consuming daily phone calls, emails, and complex electronic medical record protocols [58].

The PEU was also highly rated (average above 4). Most of the surveyed physicians (97%) declared that they know how to connect to external systems during teleconsultations. Using teleconsultations was understandable for most of the respondents (96%), and most of them (94%) could easily prepare all necessary documents (prescriptions, sick leave, referrals for tests, etc) during the teleconsultations. Our results confirm those of other studies according to which the teleconsultations are simple and support physicians' responsibility in their work and medical decisions [30,59].

Polish physicians also positively assessed the future of teleconsultations and declared their intent to this form of work with patients (83%) and other physicians to agree on the diagnosis (73%). According to the majority of respondents (79%), remote monitoring of patients' health would improve the performance of teleconsultations. In the future, they (75%) would also willingly use video visits to facilitate contact and diagnosis of patients. Similar findings were obtained in a Romanian study, in which physicians concluded that telemedicine should be used continuously, not just during the COVID-19 pandemic. Most physicians (91.1%) considered it necessary to provide care using telemedicine after the pandemic [57]. In addition, in Brazil, most physicians want to continue remote care and demand regulations on the use of telemedicine that would allow the extension of remote services [60].

Teleconsultation became popular during the COVID-19 pandemic, and now, there are expectations that it will become a permanent part of the health care system [61]. The development and integration of ICT in health care delivery have great potential for patients, providers, and payers in future health care systems [62].

XSL•FO
RenderX

Polish physicians positively assessed SAT and felt comfortable giving the system a high SAT score. The mean value of responses for all statements regarding SAT was approximately 4. Only the statement regarding the identity of remote and traditional visits was rated lower. Only 62% of physicians believed that both forms of medical consultations are equivalent, and the average for this dimension was 3.55. This is probably due to the influence of teleconsultations on Q, which Polish physicians assessed as the lowest of all the dimensions of TAM. The average response ranged from 3.28 to 3.73. Only 55.1% of respondents stated that teleconsultation is suitable for holistic patient care, and 36% stated that teleconsultation does not allow for comprehensive care. The literature also emphasizes that teleconsultations will never replace F2F meetings. The large-scale and urgent introduction of teleconsultation into our practice is likely to be redefined in the post–COVID-19 era [7]. Another opinion is that teleconsultation is not inferior to personal visits to the office in terms of the preferences and satisfaction of patients and physicians. It should, therefore, be an effective complement to F2F office visits as a mechanism for segregation and long-term continuity of care [63].

Comparison With Previous Studies

This is the first such study conducted on SAT and Q in PHCs. SAT and Q have already been studied in other medical specialties (eg, urology, dermatology, psychiatry, and oncology). In a study conducted among dermatologists, almost all categories regarding SAT with remote dermatological teleconsultations were rated at about 9 on a 0-10 scale [64]. Schubert et al [65] assessed SAT with teleconsultations in psychiatry, which turned out to be at a high level. Providers were satisfied with telepsychiatry, and both believed that telepsychiatry provides patients with better access to care. Urological teleconsultation introduced quickly during the COVID-19 closure has achieved a high level of satisfaction among both patients and physicians [7]. Physicians are interested in using telemedicine tools that increase improved access to health and differentiate their clinical practice [66]. Telemedicine benefits all physicians' patients by increasing access to health care services and remotely managing elderly people with chronic conditions [67]. Therefore, the findings of our research are in accordance with other studies documenting the openness of physicians to the use of teleconsultations in providing health services to patients [68].

However, teleconsultations have limitations regarding the uncertainty caused by the inability to physically check the patient's health condition, and this is something physicians should be aware of [69]. Studies so far show that almost two-thirds of physicians report uncertainty about the correctness of a diagnosis made with telemedicine, and only one-fourth have confidence in making remote decisions [57]. Teleconsultations will never fully replace a personal visit, due to the inability to check the physical symptoms of the disease and the lack of nonverbal signals expressing trust and empathy during remote contact.

SEM results substantiate the significant influence of PU on INT, SAT with this technology, and Q. A possible reason for this may be the availability and effectiveness of teleconsultation, the time saving in this system, and Q. Thus, a positive effect from PU will result in better SAT with teleconsultations and INT. This result is in line with other studies conducted in the field of telemedicine [30,70-73].

The medical PEU from the point of view of handling medical data (PEU_2) has a minor but significant impact on INT. Notably, the technical PEU_1 has a significant impact on Q. The easier it is to use teleconsultation, the better physicians are at assessing Q. The less effort users put in to handle medical data, the more positive their INT to use the system. This is in line with other studies, showing that the acceptance of telemedicine is greater when it provides faster health care, cost savings, better documentation, and time savings [52]. However, a study conducted in the United States found that the role of the influence of PEU on INT is insignificant. The study focused on pediatricians' INT to use of online health apps. The reason for this could be the longer contact of physicians with telemedicine technology [64]. Another explanation is that for highly competent physicians, the effect of PEU on INT is of little importance [74].

PU and PEU are considered the main determinants that directly explain the intent to use ("accept") a new technology [75]. In this study, we, therefore, confirmed the hypotheses that the constructs described in the traditional TAM are appropriate for measuring the intent of physicians to use teleconsultations.

INT emerged as a factor influencing SAT and Q and a pivotal mediator linking PU with both SAT and Q, thus underscoring the importance of intentions in the acceptance and effective use of telemedicine. This finding is in harmony with the existing literature that emphasizes the mediating role of INT in the context of technology acceptance [26]. When physicians believe that using teleconsultation will be effortless and useful, their attitude and INT will improve. This system, with less effort, encourages physicians to use teleconsultations and improves SAT and Q [31,70,76,77]. Therefore, the condition for the implementation of telemedicine technologies should be ensuring its understanding by health care providers in order to gain their acceptance and ensure the use of these technologies in the future [78]. SAT also turned out to be a significant mediator between PU, PEU, and Q. Our research, therefore, showed that the main elements of TAM viewed as PU and PEU have a significant impact on INT, which has been confirmed in other studies [30,70-73,79-82].

The model reaffirms the significant direct and indirect roles of PU and PEU in shaping INT, SAT, and Q. These findings contribute to the extant literature on teleconsultation acceptance and underscore the nuanced factors that influence the acceptance and satisfaction of teleconsultation services among physicians.

Implications

When planning a new teleconsultation system, PHC facilities should be able to predict whether the new system will be acceptable and satisfactory for medical staff, investigate the reasons why the planned system may not be fully acceptable, and then take action to increase the system's acceptance. The results of this study show that the PU of a system is a key determinant of medical professionals' INT. Therefore, before introducing a new system to PHC facilities, managers of these

XSL•FO
RenderX

facilities can increase the acceptance of the system by involving medical personnel in the implementation process, assessing the medical personnel's perception of the system (PU and PEU) and taking appropriate actions based on this assessment. Training should also be provided to medical staff to highlight the effectiveness and usefulness of teleconsultation in PHCs. Information and training sessions should primarily focus on how teleconsultation can help improve the quality of PHCs.

Intention as an intermediary variable has a significant and positive impact on users' SAT with teleconsultation and its Q. To increase the expected results in the Q of teleconsultations, the teleconsultation system should be useful for health checking, improving the quality of life, and increasing the capacity for self-care. To increase PEU, the teleconsultation system should be clear, understandable, easy to learn, easy to implement, and easy to perform health checks with. To increase the perceived utility, the teleconsultation system should positively influence the treatment plan, provide more comprehensive care services, and efficiently diagnose and efficiently plan and precisely monitor the patient's condition. The system should make physicians willing to use it to increase their INT. All this will contribute to greater SAT of physicians with their work and better quality of care [83].

This study is a contribution to the field of teleconsultation acceptance research. The modified TAM and its psychometric properties verified in this study can be used as a research framework to understand the acceptance of teleconsultation, especially in the population of PHC workers. The model can also be used for future TAM research in a variety of contexts in identifying, explaining, and predicting the intention of PHC professionals to use teleconsultation. Therefore, it is highly recommended to replicate this study in different environments to generalize the results across domains.

Limitations

In the discussion of research findings, it is crucial to acknowledge the limitations of this study to provide a comprehensive understanding of its context and implications. Although the study offers valuable insights into the factors influencing the acceptance and SAT of medical professionals with teleconsultation systems, several limitations must be considered.

The study was conducted across a specific number of PHC facilities in Poland. Although efforts were made to ensure a representative sample of PHC facilities, the findings might not fully encapsulate the diverse range of experiences and perceptions of all medical professionals nationwide. Regional variations, different health care settings, and varying levels of teleconsultation acceptance could influence SAT and acceptance levels.

The cross-sectional nature of the survey limits our ability to infer causality or changes over time. Longitudinal studies would be required to understand how perceptions and SAT with teleconsultation evolve as users gain more experience and as the technology itself advances.

The reliance on self-reported data can introduce biases, such as social desirability or recall bias. Participants' responses might not accurately reflect their true experiences or feelings toward teleconsultation.

Although the study focused on PU and PEU, other factors could influence SAT and INT. These might include individual technological proficiency, prior experiences with teleconsultation, or organizational support, which were not extensively explored in this study.

The study primarily addressed teleconsultation in PHC settings. The findings might not be generalizable to other forms of teleconsultation or to specialists' use of teleconsultation, where different factors could be more influential. The study also did not deeply explore the technological and operational constraints that might impact the effectiveness and user SAT with teleconsultation systems, such as system reliability, user interface design, and integration with existing health information systems.

The study was conducted during a period potentially influenced by the COVID-19 pandemic, which might have affected attitudes toward teleconsultation. The urgency and necessity of teleconsultation during the pandemic might not reflect standard operational conditions.

By addressing these limitations, future research can build upon the findings to develop a more nuanced understanding of the factors influencing the successful implementation and adoption of teleconsultation systems in various health care settings. In the forthcoming models, we also intend to include constructs such as compatibility, self-efficacy, social norms, perceived behavioral control [25,84], social interaction, invasiveness, and relevance [85].

Conclusion

After the outbreak of the COVID-19 pandemic, there was a dynamic development of teleconsultations in PHCs in Poland. Therefore, we conducted satisfaction surveys of Polish physicians based on a modified TAM, which we extended with new constructs, including physicians' SAT and Q, considering INT as a mediating variable. The tool developed for this model was verified in terms of psychometric properties. Therefore, it has the potential to be used in both research and practice, especially to assess the SAT and Q of PHC physicians who use teleconsultations in Poland.

The findings highlight significant relationships between PEU, PU, INT, and physicians' SAT with teleconsultation and their Q. The study showed that the PEU and PU of teleconsultations are predictive determinants of the acceptance of teleconsultation, which in turn influences physicians' SAT and Q.

Identification of the most important factors influencing physicians' SAT and Q can provide important information to managers of PHC facilities and help them make the right decisions. This study provides information for the strategies of PHCs and policy makers to accept and encourage the use of teleconsultations in Poland.

XSL•FO
RenderX

Acknowledgments

This research was supported by the Warsaw University of Technology under the Excellence Initiative: Research University (IDUB) program.

Data Availability

The data sets generated and analyzed during this study are available from the corresponding author upon reasonable request.

Authors' Contributions

RW, MK-A, and LH contributed to study conception and design and data collection and wrote the draft manuscript. RW contributed to the formal analysis and created the figures. All authors have reviewed the results and approved the final version of the manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Questionnaire statements used in the modified Technology Acceptance Model.
[DOCX File , 17 KB - humanfactors_v11i1e47810_app1.docx]

References

1. Hu PJ, Chau PY, Sheng ORL, Tam KY. Examining the technology acceptance model using physician acceptance of telemedicine technology. J Manag Inf Syst 2015 Dec 02;16(2):91-112. [doi: 10.1080/07421222.1999.11518247]
2. Walczak R, Kludacz-Alessandri M, Hawrysz L. Use of telemedicine technology among general practitioners during COVID-19: a modified technology acceptance model study in Poland. Int J Environ Res Public Health 2022 Sep 01;19(17):10937 [FREE Full text] [doi: 10.3390/ijerph191710937] [Medline: 36078650]
3. Kalaivani K, Anjalipriya V, Sivakumar R, Srimeena R. An efficient bio-key management scheme for telemedicine applications. 2015 Jul Presented at: 2015 IEEE Technological Innovation in ICT for Agriculture and Rural Development (TIAR); July 10-12, 2015; Chennai, India. [doi: 10.1109/TIAR.2015.7358543]
4. Kludacz-Alessandri M, Hawrysz L, Korneta P, Gierszewska G, Pomaranik W, Walczak R. The impact of medical teleconsultations on general practitioner-patient communication during COVID- 19: a case study from Poland. PLoS One 2021 Jul 16;16(7):e0254960 [FREE Full text] [doi: 10.1371/journal.pone.0254960] [Medline: 34270587]
5. Mann DM, Chen J, Chunara R, Testa PA, Nov O. COVID-19 transforms health care through telemedicine: evidence from the field. J Am Med Inform Assoc 2020 Jul 01;27(7):1132-1135 [FREE Full text] [doi: 10.1093/jamia/ocaa072] [Medline: 32324855]
6. Liu L, Gu J, Shao F, Liang X, Yue L, Cheng Q, et al. Application and preliminary outcomes of remote diagnosis and treatment during the COVID-19 outbreak: retrospective cohort study. JMIR Mhealth Uhealth 2020 Jul 03;8(7):e19417 [FREE Full text] [doi: 10.2196/19417] [Medline: 32568722]
7. Pinar U, Anract J, Perrot O, Tabourin T, Chartier-Kastler E, Parra J, et al. Preliminary assessment of patient and physician satisfaction with the use of teleconsultation in urology during the COVID-19 pandemic. World J Urol 2021 Jun 09;39(6):1991-1996 [FREE Full text] [doi: 10.1007/s00345-020-03432-4] [Medline: 32909174]
8. Layfield E, Triantafillou V, Prasad A, Deng J, Shanti RM, Newman JG, et al. Telemedicine for head and neck ambulatory visits during COVID-19: evaluating usability and patient satisfaction. Head Neck 2020 Jul;42(7):1681-1689 [FREE Full text] [doi: 10.1002/hed.26285] [Medline: 32476228]
9. Tenforde AS, Borgstrom H, Polich G, Steere H, Davis IS, Cotton K, et al. Outpatient physical, occupational, and speech therapy synchronous telemedicine: . a survey study of patient satisfaction with virtual visits during the COVID-19 pandemic. Am J Phys Med Rehabil 2020 Nov;99(11):977-981 [FREE Full text] [doi: 10.1097/PHM.0000000000001571] [Medline: 32804713]
10. Bhattacherjee A, Hikmet N. Physicians' resistance toward healthcare information technology: a theoretical model and empirical test. Eur J Inf Syst 2017 Dec 19;16(6):725-737. [doi: 10.1057/palgrave.ejis.3000717]
11. Davis F. Usefulness, perceived ease of use, and user acceptance of information technology. MIS Quarterly 1989;13(3):319-340 [FREE Full text] [doi: 10.2307/249008]
12. Venkatesh V, Morris MG, Davis GB, Davis FD. User acceptance of information technology: toward a unified view. MIS Quarterly 2003 Sep;27(3):425-478. [doi: 10.2307/30036540]
13. Venkatesh V, Davis FD. A theoretical extension of the technology acceptance model: four longitudinal field studies. Manag Sci 2000 Feb;46(2):186-204. [doi: 10.1287/mnsc.46.2.186.11926]
14. Beldad AD, Hegner SM. Expanding the technology acceptance model with the inclusion of trust, social influence, and health valuation to determine the predictors of German users' willingness to continue using a fitness app: a structural equation modeling approach. Int J Hum–Comput Interact 2017 Nov 30;34(9):882-893. [doi: 10.1080/10447318.2017.1403220]

XSL•FO
RenderX

15. Legris P, Ingham J, Collerette P. Why do people use information technology? A critical review of the technology acceptance model. Inf Manag 2003 Jan;40(3):191-204. [doi: 10.1016/S0378-7206(01)00143-4]
16. Yarbrough AK, Smith TB. Technology acceptance among physicians: a new take on TAM. Med Care Res Rev 2007 Dec;64(6):650-672. [doi: 10.1177/1077558707305942] [Medline: 17717378]
17. Huang J. Innovative health care delivery system--a questionnaire survey to evaluate the influence of behavioral factors on individuals' acceptance of telecare. Comput Biol Med 2013 May;43(4):281-286. [doi: 10.1016/j.compbiomed.2012.12.011] [Medline: 23375377]
18. Rho MJ, Choi IY, Lee J. Predictive factors of telemedicine service acceptance and behavioral intention of physicians. Int J Med Inform 2014 Aug;83(8):559-571. [doi: 10.1016/j.ijmedinf.2014.05.005] [Medline: 24961820]
19. Ladan MA, Wharrad H, Windle R. Towards understanding healthcare professionals' adoption and use of technologies in clinical practice: using Qmethodology and models of technology acceptance. J Innov Health Inform 2018 Mar 09;25(1):965-937 [FREE Full text] [doi: 10.14236/jhi.v25i1.965] [Medline: 29717952]
20. Van Schaik P, Bettany-Saltikov JA, Warren JG. Clinical acceptance of a low-cost portable system for postural assessment. Behav Inf Technol 2002 Jan;21(1):47-57. [doi: 10.1080/01449290110107236]
21. Chau PYK, Hu PJ. Information technology acceptance by individual professionals: a model comparison approach. Decis Sci 2007 Jun 07;32(4):699-719. [doi: 10.1111/j.1540-5915.2001.tb00978.x]
22. Holden RJ, Karsh B. The technology acceptance model: its past and its future in health care. J Biomed Inform 2010 Feb;43(1):159-172 [FREE Full text] [doi: 10.1016/j.jbi.2009.07.002] [Medline: 19615467]
23. Croteau A, Vieru D. Telemedicine adoption by different groups of physicians. 2002 Presented at: 35th Annual Hawaii International Conference on System Sciences; January 7-10, 2002; Big Island, HI p. 1985-1993. [doi: 10.1109/hicss.2002.994121]
24. Hsiao J, Chen R. Critical factors influencing physicians' intention to use computerized clinical practice guidelines: an integrative model of activity theory and the technology acceptance model. BMC Med Inform Decis Mak 2016 Jan 16;16(1):3 [FREE Full text] [doi: 10.1186/s12911-016-0241-3] [Medline: 26772169]
25. Deng Q, Zheng Y, Lu J, Zeng Z, Liu W. What factors predict physicians' utilization behavior of contrast-enhanced ultrasound? Evidence from the integration of the Theory of Planned Behavior and Technology Acceptance Model using a structural equation modeling approach. BMC Med Inform Decis Mak 2021 May 31;21(1):173 [FREE Full text] [doi: 10.1186/s12911-021-01540-8] [Medline: 34059025]
26. Kissi J, Dai B, Dogbe CS, Banahene J, Ernest O. Predictive factors of physicians' satisfaction with telemedicine services acceptance. Health Informat J 2020 Sep 19;26(3):1866-1880 [FREE Full text] [doi: 10.1177/1460458219892162] [Medline: 31854222]
27. Metallo C, Agrifoglio R, Lepore L, Landriani L. Explaing users' technology acceptance through national cultural values in the hospital context. BMC Health Serv Res 2022 Jan 17;22(1):84 [FREE Full text] [doi: 10.1186/s12913-022-07488-3] [Medline: 35039014]
28. Lin Y, Yu C, Yi C. The effects of positive affect, person-job fit, and well-being on job performance. Soc Behav Pers 2014 Oct 26;42(9):1537-1547. [doi: 10.2224/sbp.2014.42.9.1537]
29. Zayyad M, Toycan M. Factors affecting sustainable adoption of e-health technology in developing countries: an exploratory survey of Nigerian hospitals from the perspective of healthcare professionals. PeerJ 2018;6:e4436 [FREE Full text] [doi: 10.7717/peerj.4436] [Medline: 29507830]
30. Chau PY, Hu PJ. Investigating healthcare professionals' decisions to accept telemedicine technology: an empirical test of competing theories. Inf Manag 2002 Jan;39(4):297-311. [doi: 10.1016/S0378-7206(01)00098-2]
31. Tubaishat A. Perceived usefulness and perceived ease of use of electronic health records among nurses: application of technology acceptance model. Inform Health Soc Care 2018 Dec;43(4):379-389. [doi: 10.1080/17538157.2017.1363761] [Medline: 28920708]
32. Vitari C, Ologeanu-Taddei R. The intention to use an electronic health record and its antecedents among three different categories of clinical staff. BMC Health Serv Res 2018 Mar 21;18(1):194 [FREE Full text] [doi: 10.1186/s12913-018-3022-0] [Medline: 29562942]
33. Yi MY, Jackson JD, Park JS, Probst JC. Understanding information technology acceptance by individual professionals: toward an integrative view. Inf Manag 2006 Apr;43(3):350-363. [doi: 10.1016/j.im.2005.08.006]
34. Premkumar G, Bhattacherjee A. Explaining information technology usage: a test of competing models. Omega 2008 Feb;36(1):64-75. [doi: 10.1016/j.omega.2005.12.002]
35. Bhattacherjee A, Premkumar G. Understanding changes in belief and attitude toward information technology usage: a theoretical model and longitudinal test. MIS Quarterly 2004;28(2):229. [doi: 10.2307/25148634]
36. Bhattacherjee A. Understanding information systems continuance: an expectation-confirmation model. MIS Quarterly 2001 Sep;25(3):351. [doi: 10.2307/3250921]
37. Alsohime F, Temsah MH, Al-Eyadhy A, Bashiri FA, Househ M, Jamal A, et al. Satisfaction and perceived usefulness with newly-implemented electronic health records system among pediatricians at a university hospital. Comput Methods Programs Biomed 2019 Feb;169:51-57. [doi: 10.1016/j.cmpb.2018.12.026] [Medline: 30638591]

XSL•FO
RenderX

38. Petter S, Fruhling A. Evaluating the success of an emergency response medical information system. Int J Med Inform 2011 Jul;80(7):480-489. [doi: 10.1016/j.ijmedinf.2011.03.010] [Medline: 21501969]
39. Padilha JM, Machado PP, Ribeiro AL, Ribeiro R, Vieira F, Costa P. Easiness, usefulness and intention to use a MOOC in nursing. Nurse Educ Today 2021 Feb;97:104705. [doi: 10.1016/j.nedt.2020.104705] [Medline: 33341525]
40. Saputra DEE, Maulida Rahma V, Eliyana A, Pratama AS, Anggraini RD, Kamil NLM, et al. Do system quality and information quality affect job performance? The mediation role of users' perceptions. PLoS One 2023;18(6):e0285293 [FREE Full text] [doi: 10.1371/journal.pone.0285293] [Medline: 37352227]
41. Chirchir LK, Aruasa WK, Chebon SK. Perceived usefulness and ease of use as mediators of the effect of health information systems on user performance. Eur J Comput Sci Inf Technol 2019 Feb;7(1):22-37. [doi: 10.1017/CBO9781107415324.004]
42. Souza-Pereira L, Ouhbi S, Pombo N. A process model for quality in use evaluation of clinical decision support systems. J Biomed Inform 2021 Nov;123:103917 [FREE Full text] [doi: 10.1016/j.jbi.2021.103917] [Medline: 34563692]
43. Korneta P, Kludacz-Alessandri M, Walczak R. The impact of COVID-19 on the performance of primary health care service providers in a capitation payment system: a case study from Poland. Int J Environ Res Public Health 2021 Feb 03;18(4):1-19 [FREE Full text] [doi: 10.3390/ijerph18041407] [Medline: 33546467]
44. Davis JP, Crombie IK, Davies HT. Understanding sampling: representativeness matters. Hosp Med 2001 Apr;62(4):237-240. [doi: 10.12968/hosp.2001.62.4.2052] [Medline: 11338957]
45. Wetterberg H, Rydén L, Ahlner F, Falk Erhag H, Gudmundsson P, Guo X, et al. Representativeness in population-based studies of older adults: five waves of cross-sectional examinations in the Gothenburg H70 Birth Cohort Study. BMJ Open 2022 Dec 16;12(12):e068165 [FREE Full text] [doi: 10.1136/bmjopen-2022-068165] [Medline: 36526314]
46. Lynn P. Maintaining cross-sectional representativeness in a longitudinal general population survey. Understanding Society Working Paper Series No. 2011 – 04. 2011 Jun. URL: https://www.understandingsociety.ac.uk/wp-content/uploads/working-papers/2011-04.pdf [accessed 2023-12-15]
47. Hoffmann R, Lange M, Butschalowsky H, Houben R, Schmich P, Allen J, et al. KiGGS wave 2 cross-sectional study - participant acquisition, response rates and representativeness. J Health Monit 2018 Mar;3(1):78-91 [FREE Full text] [doi: 10.17886/RKI-GBE-2018-032] [Medline: 35586176]
48. Martínez A, Everss E, Rojo-Alvarez JL, Figal DP, García-Alberola A. A systematic review of the literature on home monitoring for patients with heart failure. J Telemed Telecare 2006;12(5):234-241. [doi: 10.1258/135763306777889109] [Medline: 16848935]
49. Rimmer MP, Al Wattar BH, UKARCOG Members. Provision of obstetrics and gynaecology services during the COVID-19 pandemic: a survey of junior doctors in the UK National Health Service. BJOG 2020 Aug;127(9):1123-1128 [FREE Full text] [doi: 10.1111/1471-0528.16313] [Medline: 32460422]
50. Bakken S, Grullon-Figueroa L, Izquierdo R, Lee N, Morin P, Palmas W, IDEATel Consortium. Development, validation, and use of English and Spanish versions of the telemedicine satisfaction and usefulness questionnaire. J Am Med Inform Assoc 2006;13(6):660-667 [FREE Full text] [doi: 10.1197/jamia.M2146] [Medline: 16929036]
51. Rho MJ, Kim SR, Kim H, Cho J, Yoon K, Mun SK, et al. Exploring the relationship among user satisfaction, compliance, and clinical outcomes of telemedicine services for glucose control. Telemed J E Health 2014 Aug;20(8):712-720. [doi: 10.1089/tmj.2013.0309] [Medline: 24901576]
52. Kamal SA, Shafiq M, Kakria P. Investigating acceptance of telemedicine services through an extended technology acceptance model (TAM). Technol Soc 2020 Feb;60:101212. [doi: 10.1016/j.techsoc.2019.101212]
53. Schepers J, Wetzels M. A meta-analysis of the technology acceptance model: investigating subjective norm and moderation effects. Inf Manag 2007 Jan;44(1):90-103. [doi: 10.1016/j.im.2006.10.007]
54. Kohnke A, Cole ML, Bush R. Incorporating UTAUT predictors for understanding home care patients' and clinician's acceptance of healthcare telemedicine equipment. J Technol Manag Innovat 2014 Jul;9(2):29-41. [doi: 10.4067/S0718-27242014000200003]
55. Hair J, Babin B, Anderson R, Black W. Multivariate Data Analysis. 8th Edition. Noida, Uttar Pradesh, India: Cengage; 2018.
56. Tavakol M, Dennick R. Making sense of Cronbach's alpha. Int J Med Educ 2011 Jun 27;2:53-55 [FREE Full text] [doi: 10.5116/ijme.4dfb.8dfd] [Medline: 28029643]
57. Florea M, Lazea C, Gaga R, Sur G, Lotrean L, Puia A, et al. Lights and shadows of the perception of the use of telemedicine by Romanian family doctors during the COVID-19 pandemic. Int J Gen Med 2021;14:1575-1587 [FREE Full text] [doi: 10.2147/IJGM.S309519] [Medline: 33953605]
58. Thornton J. Covid-19: how coronavirus will change the face of general practice forever. BMJ 2020 Mar 30;368:m1279. [doi: 10.1136/bmj.m1279] [Medline: 32229477]
59. Craig J, Patterson V. Introduction to the practice of telemedicine. J Telemed Telecare 2005 Jun;11(1):3-9. [doi: 10.1177/1357633X0501100102] [Medline: 15829036]
60. Caetano R, Silva AB, Guedes ACCM, Paiva CCND, Ribeiro GDR, Santos DL, et al. Challenges and opportunities for telehealth during the COVID-19 pandemic: ideas on spaces and initiatives in the Brazilian context. Cad Saude Publica 2020;36(5):e00088920 [FREE Full text] [doi: 10.1590/0102-311x00088920] [Medline: 32490913]

XSL•FO
RenderX

61. Hoff T, Lee DR. Physician satisfaction with telehealth: a systematic review and agenda for future research. Qual Manag Health Care 2022;31(3):160-169 [FREE Full text] [doi: 10.1097/QMH.0000000000000359] [Medline: 35132008]
62. Abdel-Wahab M, Rosenblatt E, Prajogi B, Zubizarretta E, Mikhail M. Opportunities in telemedicine, lessons learned after COVID-19 and the way into the future. Int J Radiat Oncol Biol Phys 2020 Oct 01;108(2):438-443 [FREE Full text] [doi: 10.1016/j.ijrobp.2020.07.006] [Medline: 32890528]
63. Melian C, Kieser D, Frampton C, C Wyatt M. Teleconsultation in orthopaedic surgery: a systematic review and meta-analysis of patient and physician experiences. J Telemed Telecare 2022 Aug 01;28(7):471-480. [doi: 10.1177/1357633X20950995] [Medline: 32873138]
64. Fluhr JW, Gueguen A, Legoupil D, Brenaut E, Abasq C, Araújo H, et al. Teledermatology in times of COVID-19 confinement: comparing patients' and physicians' satisfaction by the standardized brest teledermatology questionnaire. Dermatology 2021 Feb 10;237(2):1-6 [FREE Full text] [doi: 10.1159/000514029] [Medline: 33567427]
65. Schubert NJ, Backman PJ, Bhatla R, Corace KM. Telepsychiatry and patient-provider concordance. Can J Rural Med 2019;24(3):75-82 [FREE Full text] [doi: 10.4103/CJRM.CJRM_9_18] [Medline: 31249155]
66. Hodgkins M, Barron M, Jevaji S, Lloyd S. Physician requirements for adoption of telehealth following the SARS-CoV-2 pandemic. NPJ Digit Med 2021 Feb 09;4(1):19 [FREE Full text] [doi: 10.1038/s41746-021-00390-y] [Medline: 33564102]
67. Florea M, Puia A, Pop R. The family as recipient and provider of home care: a primary care perspective. In: Cascella M, Stones MJ, editors. Suggestions for Addressing Clinical and Non-Clinical Issues in Palliative Care. London, UK: IntechOpen; 2021.
68. Helou S, El Helou E, Abou-Khalil V, Wakim J, El Helou J, Daher A, et al. The effect of the COVID-19 pandemic on physicians' use and perception of telehealth: the case of Lebanon. Int J Environ Res Public Health 2020 Jul 06;17(13):1-17 [FREE Full text] [doi: 10.3390/ijerph17134866] [Medline: 32640652]
69. Iyengar K, Jain VK, Vaishya R. Pitfalls in telemedicine consultations in the era of COVID 19 and how to avoid them. Diabetes Metab Syndr 2020;14(5):797-799 [FREE Full text] [doi: 10.1016/j.dsx.2020.06.007] [Medline: 32534432]
70. Abdullah F, Ward R, Ahmed E. Investigating the influence of the most commonly used external variables of TAM on students' perceived ease of use (PEOU) and perceived usefulness (PU) of e-portfolios. Comput Hum Behav 2016 Oct;63:75-90. [doi: 10.1016/j.chb.2016.05.014]
71. Marangunić N, Granić A. Technology acceptance model: a literature review from 1986 to 2013. Univ Access Inf Soc 2014 Feb 16;14(1):81-95. [doi: 10.1007/s10209-014-0348-1]
72. Khatun F, Heywood AE, Ray PK, Hanifi SMA, Bhuiya A, Liaw S. Determinants of readiness to adopt mHealth in a rural community of Bangladesh. Int J Med Inform 2015 Oct;84(10):847-856. [doi: 10.1016/j.ijmedinf.2015.06.008] [Medline: 26194141]
73. Venkatesh V, Sykes T, Zhang X. 'Just what the doctor ordered': a revised UTAUT for EMR system adoption and use by doctors. 2021 Presented at: 2011 44th Hawaii International Conference on System Sciences; January 4-7, 2011; Kauai, HI. [doi: 10.1109/hicss.2011.1]
74. Chismar WG, Wiley-Patton S. Test of the technology acceptance model for the internet in pediatrics. Proc AMIA Symp 2002:155-159 [FREE Full text] [Medline: 12463806]
75. Scherer R, Siddiq F, Tondeur J. The technology acceptance model (TAM): a meta-analytic structural equation modeling approach to explaining teachers' adoption of digital technology in education. Comput Educ 2019 Jan;128:13-35. [doi: 10.1016/j.compedu.2018.09.009]
76. Hoque MR, Bao Y, Sorwar G. Investigating factors influencing the adoption of e-Health in developing countries: a patient's perspective. Inform Health Soc Care 2017 Jan;42(1):1-17. [doi: 10.3109/17538157.2015.1075541] [Medline: 26865037]
77. Wai ISH, Ng SSY, Chiu DKW, Ho KKW, Lo P. Exploring undergraduate students' usage pattern of mobile apps for education. J Librariansh Inf Sci 2016 Aug 30;50(1):34-47 [FREE Full text] [doi: 10.1177/0961000616662699]
78. Kalayou MH, Endehabtu BF, Tilahun B. The applicability of the modified technology acceptance model (TAM) on the sustainable adoption of eHealth systems in resource-limited settings. J Multidiscip Healthc 2020;13:1827-1837 [FREE Full text] [doi: 10.2147/JMDH.S284973] [Medline: 33299320]
79. Faber S, van Geenhuizen M, de Reuver M. eHealth adoption factors in medical hospitals: a focus on the Netherlands. Int J Med Inform 2017 Apr;100:77-89. [doi: 10.1016/j.ijmedinf.2017.01.009] [Medline: 28241940]
80. Melas CD, Zampetakis LA, Dimopoulou A, Moustakis V. Modeling the acceptance of clinical information systems among hospital medical staff: an extended TAM model. J Biomed Inform 2011 Aug;44(4):553-564 [FREE Full text] [doi: 10.1016/j.jbi.2011.01.009] [Medline: 21292029]
81. Khan I, Xitong G, Ahmad Z, Shahzad F. Investigating factors impelling the adoption of e-Health: a perspective of African expats in China. SAGE Open 2019 Jul 23;9(3):215824401986580 [FREE Full text] [doi: 10.1177/2158244019865803]
82. Briz-Ponce L, García-Peñalvo FJ. An empirical assessment of a technology acceptance model for apps in medical education. J Med Syst 2015 Nov;39(11):176. [doi: 10.1007/s10916-015-0352-x] [Medline: 26411928]
83. Andriani A, Pink BM. Acceptance of Halodoc's online teleconsultation during covid-19. Manag Sci Field 2022;12(2):1566-1574 [FREE Full text] [doi: 10.1016/b978-0-323-99878-9.09997-0]
84. Garavand A, Aslani N, Nadri H, Abedini S, Dehghan S. Acceptance of telemedicine technology among physicians: a systematic review. Informat Med Unlock 2022;30:100943. [doi: 10.1016/j.imu.2022.100943]

XSL•FO
RenderX

85. Zhao J, Wang J. Health advertising on short-video social media: a study on user attitudes based on the extended technology acceptance model. Int J Environ Res Public Health 2020 Feb 26;17(5):1501 [FREE Full text] [doi: 10.3390/ijerph17051501] [Medline: 32110950]

Abbreviations

CR: critical ratio
EFA: exploratory factor analysis
F2F: face-to-face
ICT: information and communication technology
INT: intention to use teleconsultation
KMO: Kaiser-Meyer-Olkin
NHF: National Health Fund
PEU: perceived ease of use
PHC: primary health care facility
PU: perceived usefulness
Q: quality of work
SAT: satisfaction
SEM: structural equation modeling
TAM: Technology Acceptance Model

Edited by A Kushniruk; submitted 02.04.23; peer-reviewed by E Kim, J Mira; comments to author 01.08.23; revised version received 18.12.23; accepted 20.03.24; published 10.06.24.

Please cite as:
Hawrysz L, Kludacz-Alessandri M, Walczak R
Predictive Factors of Physicians' Satisfaction and Quality of Work Under Teleconsultation Conditions: Structural Equation Analysis
JMIR Hum Factors 2024;11:e47810
URL: https://humanfactors.jmir.org/2024/1/e47810
doi:10.2196/47810
PMID:38857081

©Liliana Hawrysz, Magdalena Kludacz-Alessandri, Renata Walczak. Originally published in JMIR Human Factors (https://humanfactors.jmir.org), 10.06.2024. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on https://humanfactors.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Safety in Teletriage by Nurses and Physicians in the United States and Israel: Narrative Review and Qualitative Study

Motti Haimi[1,2,3], MD, MHA, PhD; Sheila Quilter Wheeler[4], BA, MSc

[1]Rappaport Faculty of Medicine, Technion – Israel Institute of Technology, Haifa, Israel
[2]Health Systems Management Department, The Max Stern Yezreel Valley College, Emek Yezreel, Israel
[3]Meuhedet Healthcare Services – North District, Tel Aviv, Israel
[4]TeleTriage Systems, San Anselmo, CA, United States

Corresponding Author:
Motti Haimi, MD, MHA, PhD
Rappaport Faculty of Medicine
Technion – Israel Institute of Technology
Efron 1, Bat Galim
Haifa, 3200003
Israel
Phone: 972 504557767
Email: morx@netvision.net.il

Abstract

Background: The safety of telemedicine in general and telephone triage (teletriage) safety in particular have been a focus of concern since the 1970s. Today, telehealth, now subsuming teletriage, has a basic structure and process intended to promote safety. However, inadequate telehealth systems may also compromise patient safety. The COVID-19 pandemic accelerated rapid but uneven telehealth growth, both technologically and professionally. Within 5-10 years, the field will likely be more technologically advanced; however, these advances may still outpace professional standards. The need for an evidence-based system is crucial and urgent.

Objective: Our aim was to explore ways that developed teletriage systems produce safe outcomes by examining key system components and questioning long-held assumptions.

Methods: We examined safety by performing a narrative review of the literature using key terms concerning patient safety in teletriage. In addition, we conducted system analysis of 2 typical formal systems, physician led and nurse led, in Israel and the United States, respectively, and evaluated those systems' respective approaches to safety. Additionally, we conducted in-depth interviews with representative physicians and 1 nurse using a qualitative approach.

Results: The review of literature indicated that research on various aspects of telehealth and teletriage safety is still sparse and of variable quality, producing conflicting and inconsistent results. Researchers, possibly unfamiliar with this complicated field, use an array of poorly defined terms and appear to design studies based on unfounded assumptions. The interviews with health care professionals demonstrated several challenges encountered during teletriage, mainly making diagnosis from a distance, treating unfamiliar patients, a stressful atmosphere, working alone, and technological difficulties. However, they reported using several measures that help them make accurate diagnoses and reasonable decisions, thus keeping patient safety, such as using their expertise and intuition, using structured protocols, and considering nonmedical factors and patient preferences (shared decision-making).

Conclusions: Remote encounters about acute, worrisome symptoms are time sensitive, requiring decision-making under conditions of uncertainty and urgency. Patient safety and safe professional practice are extremely important in the field of teletriage, which has a high potential for error. This underregulated subspecialty lacks adequate development and substantive research on system safety. Research may commingle terminology and widely different, ill-defined groups of decision makers with wide variation in decision-making skills, clinical training, experience, and job qualifications, thereby confounding results. The rapid pace of telehealth's technological growth creates urgency in identifying safe systems to guide developers and clinicians about needed improvements.

(JMIR Hum Factors 2024;11:e50676) doi:10.2196/50676

XSL•FO
RenderX

KEYWORDS

telephone triage; teletriage; telehealth; telemedicine; safety; system error; human error; triage; outcome; patient safety

Introduction

Definitions and Terminology

Telemedicine refers to the electronic transmission of medical data from one source to another to promote clinical health. A wide range of services and applications, including 2-way video, email, smartphones, and other communications technology, are included in telemedicine. With the aid of these technologies, patients and caregivers who are geographically separated can communicate and receive treatment, consultation, follow-up, counseling, and health education, as well as engage in medical intervention, monitoring, and remote hospitalization [1,2]. The biggest benefits of telemedicine, aside from cost savings, are expanding patient access to treatment, expanding the availability of medical services, and improving clinicians' efficacy [3].

The delivery and facilitation of health and health-related services, such as medical care, provider and patient education, health information services, and self-care, using telecommunications and digital communication technology is known as *telehealth*.

Although telehealth and telemedicine are frequently used synonymously, the term "telehealth" is used as an umbrella term to refer to all aspects and activities of health care and the health care system that are carried out via telecommunications technology, as opposed to the more specific term "telemedicine," which only refers to the practice of medicine remotely [4,5].

Triage is the process of classifying and prioritizing symptoms. Based on quality, and in the context of health services, triage refers to the process of ranking patients according to their need for care. Using triage services can lower health care costs by preventing patients from making needless and expensive trips to emergency departments (EDs) and by assisting them with self-care and informal care while the doctor is away [6,7].

Telephone triage (assessment and triage of symptoms by telephone) predates telehealth by about 50 years. In the past 5-10 years, the broad industry of telehealth has subsumed telephone triage, which has quickly evolved into *teletriage* to include a wide range of high-tech features (video, biotelemetry, and patient wearables) to enhance remote, brief, but urgent encounters [6,8].

Teletriage is an unscheduled, brief (2-10 minutes), urgent encounter (by telephone only) initiated by patients seeking an estimate of symptom urgency and triage by a clinician to get an urgent on-site evaluation and definitive diagnosis [6,8].

Televisits (via video technology) are now a common substitute for a face-to-face medical appointment and may be 20-30 minutes in length.

Definitions and terms related to telehealth and teletriage are included in Multimedia Appendix 1.

Teletriage: History and Characteristics

Wheeler et al [8] defined teletriage as the complex process of remotely assessing acute, worrisome symptoms to estimate their urgency and to render clinical advice and triage for further evaluation and diagnosis, as appropriate. The goal is to ensure the safe, timely, and appropriate disposition of patient symptoms remotely. This service is accomplished with remote encounters by telephone or real-time video (including biotelemetry). A disposition is a directive from clinician to patient about when and where to be further evaluated and treated. It may also include a risk estimation statement, such as "your symptoms sound urgent," to both inform and motivate patients to comply with the medical advice [8].

Historically, the need for teletriage became an issue when health maintenance organizations (HMOs) realized that they could be more cost-effective by conserving on-site appointments for the sickest patients, which is a form of triage and a way to control access [9].

The overarching goal of medical care (systems and processes) is to use valid, reliable components and experienced trained clinicians to produce safe outcomes. Since the 1970s, clinicians have informally performed teletriage in ambulatory care settings ranging from urgent care and EDs, physician offices, clinics, and student health centers to disease management and ambulatory surgery. Beginning in the midnineties, teletriage and the telehealth industry began developing early systems [10].

One description of teletriage [10] is that it is a time-sensitive, complex, human-technology hybrid process of remote medical decision-making. Currently and in the future, a range of technologies will provide a range of information. On the continuum of care, teletriage can now be acknowledged as the entry point to clinical care. It legitimately qualifies as "prehospital care."

By discussing treatment alternatives and the need for care, teletriage aims to identify the most appropriate degree of care that is needed. These alternatives could involve self-care or informal care, normal or emergency doctor visits, prompt referral to the ED or emergency clinics, or ambulance dispatch, depending on the data collected during evaluations. To support self-care and informal care, teletriage services may also entail providing information and help for difficult medical decisions, as well as managing symptoms. In a variety of medical facilities, nurses or doctors conduct teletriage [11,12].

The teletriage system, which is primarily run by nurses, determines the level of medical urgency and the type of health care that is needed when patients are contacted by telephone. This system is crucial for delivering affordable, effective, and secure health care [13].

The decision-making process is difficult and stressful in the emergency teletriage scenario because decisions must be made quickly and are dependent on nonvisible, unreliable, and incomplete information and nonvisible indicators. Additionally, patients' capacities for clear symptom communication differ,

XSL•FO
RenderX

particularly when young patients are involved. The lack of precise criteria for making the decision further increases the difficulty of the decision-making process [14,15].

Over the past few years, numerous Western nations and sizable corporations have started to offer primary health care services after regular business hours. In 2020, the COVID-19 pandemic accelerated telehealth growth exponentially. Almost overnight, telehealth rapidly became an established, essential service [16].

Currently, many US health plans provide advice lines. These services are advertised as a benefit of patients' health plans. Advice lines offer clinicians' advice for patients who have concerns about acute or worrisome symptoms who are calling from home. A telephone call to an after-hours advice line is typically the patients' first attempt to gain access—a medical consultation for a symptom that patients interpret as urgent. However, standards are still lacking for clinical decision makers, their experience, qualifications, clinical training, and practice [8,17,18].

Once considered embryonic, telehealth now appears to be in an adolescent phase. It is rapidly and erratically growing, and technology is outpacing clinical standards. Telehealth appears to be undergoing an identity crisis [9].

Opposing forces—technology, cost containment, and safe clinical practice—now struggle to claim control of the field, one so new that regulation cannot keep pace. There are inherent risks in the clinical task—remote, rapid, clinical decision-making using software that serves technological interests but may not serve clinical safety [18].

In the United States, these forces are quite evident, as health care needs to save money may be at odds with patients' needs for access. Furthermore, health care institutions may attempt to limit patients' use of ED, urgent care, office, and clinic services to be cost-effective or to use less costly (and less qualified) staff in the telehealth process, thereby reducing safety [10,19].

In the United States, evidence-based electronic guidelines have not yet emerged. No telehealth-based professional organization yet exists. Some agencies have developed regulations, including the Utilization Review Accreditation Commission (URAC) [20], the American Academy of Ambulatory Care Nursing [21], the American College of Emergency Physicians (ACEP) [22], the American Nurses Association (ANA) [23], the Emergency Nurses Association [24], and the North American Nursing Diagnosis Association (NANDA) [25]. Inadequately designed technology can lead to unintended consequences, while field testing may not be adequate [26,27]. The ACEP [22] has developed descriptions and broad classifications of emergent to acute symptom patterns for on-site triage. There is a need for a similar classification for teletriage.

Clearly, we are in early days in telehealth research, with the need to define meaningful measures for safe outcomes (Table 1).

Table 1. Safe outcomes.

Outcome	Description
AR[a]	A timely, safe disposition: right place, right time, for the right reason. ARs are considered safe.
OR[b]	A referral deemed (retrospectively) by some to be unnecessary at the time and place initially recommended. ORs are likely safe but not cost-effective.
UR[c]	A referral to a lower level of care than is safe or timely, often resulting in a delay in care. URs have the potential to cause or result in patient harm [8].

[a]AR: appropriate referral.

[b]OR: overreferral.

[c]UR: underreferral.

Controversies have emerged in relation to referrals (outcomes). Both appropriate referrals (ARs) and overreferrals (ORs) are considered safe, but ORs are not cost-effective—a less desirable outcome. Some experts suspect that doctors of medicine (MDs) are reluctant to define safety or to criticize other MDs'/researchers' work.

Without a consensus on safe outcomes that are evidence based, it will be difficult for the industry to make meaningful progress toward the goal of safety. Research on meaningful safe outcomes is needed. We chose to discuss teletriage safety for several reasons: patient safety and safe practice are important topics, and teletriage has a high potential for error.

Teletriage involves making medical decisions under conditions of uncertainty and urgency. Teletriage has also conflicting goals: the institutions' need to control costs (especially in the United States), while also ensuring patient access to care in a safe, timely manner. Furthermore, this underregulated subspecialty lacks adequate research on system safety.

Our review and analysis present a glimpse of current safety through analyzing 2 developed representative systems in 2 countries: Israel (Clalit Health Services) and the United States (Redwood Healthcare Plan [RHP]). We examined each system to learn how developed each system is and to explore the elements that might influence safe practice and patient outcomes.

This might be the first study to review and compare 2 formal teletriage systems. Both authors have performed triage in formal systems and taught and consulted in telehealth for a combined 50 years. Teletriage was the focus of this study. We focused on urgent, time-sensitive calls from patients or their families regarding acute or worrisome symptoms. We believe the telehealth industry and telemedicine can benefit from our findings.

XSL•FO
RenderX

Methods

Study Design

This study included 2 parts:

- A narrative review of the literature: studies describing nurses' and physicians' teletriage systems from the United States and Israel
- Qualitative assessment, including interviews with several physicians (Israel) and 1 nurse (the United States)

Part 1: Narrative Review

Both key authors of this paper have practiced in the field of telehealth; thus, they have a reality-based perspective on the subspecialty.

We restricted our review to system features (structure, process, and outcomes) to provide a more orderly review in this variable and broad field. In this narrative review, we discussed the various facets of and challenges in teletriage, with special focus on the United States and Israel, which serve as representatives for teletriage for nurses and physicians, respectively.

Search Terms

Using the following key search terms, we searched PubMed, Medline, and Google Scholar for papers relevant to this review: Telephone Triage AND Teletriage AND Telehealth AND Telemedicine AND Telecare (and Tele-Triage); + Safety, + Systems, + Physician-led, + Nurse-led. + System Error, + Human Error.

Selection Criteria

It was essential that we study developed systems, because even today, US telehealth practice is still typically unregulated—variously devoid of complete or evidence-based components, such as guidelines, formal documentation, qualified staff, clinical training, and standards—in many office, clinic, ED, and ambulatory care settings.

For our critical analysis, we focused on the best examples of current practice: large, formal clinical call centers or HMOs. These services in the United States provide 24–7 clinical call coverage.

We narrowed our review of the literature to studies that focused either directly or generally on teletriage safety. See Multimedia Appendix 2 for criteria for selecting papers. Only English language publications that were published in scholarly journals or organizations between 1970 and 2023 were included. All types of papers were considered, including original papers, reports of randomized clinical trials, observational studies, and editorials or essays by key opinion leaders.

A summary of early research (1977 to the 1990s) focused mostly on the physician practice of teletriage. A recent review (the 1990s to 2022) summarized and critiqued teletriage safety research.

Part 2: Qualitative Assessment

Using a qualitative approach, we conducted interviews with 15 representative physicians who worked in a pediatric teletrige service (Clalit Health Services) in Israel. In addition, we interviewed 1 nurse who worked in a nurse teltriage service in the United States.

To obtain their subjective perspectives on maintaining patient safety in this setting, the physicians were asked about factors that may have impacted their reaching a "correct" diagnosis and deciding on reasonable and appropriate treatment.

To gather detailed and accurate information that would accurately reflect the participants' subjective experiences, we used a semistructured qualitative study (SSQS) technique in this study. Participants' replies were evaluated and analyzed thematically when themes were found.

The use of open-ended questions, which gave the study its qualitative quality, allowed participants to candidly discuss the challenges they encounter in teletriage settings and the strategies they use to ensure patient safety.

The research complied with the Standards for Qualitative Research (SRQR) items [28]. We examined the responses using qualitative content analysis, which is a systematic procedure for collecting and analyzing qualitative data. Using a consistent set of codes to group texts with comparable content and creating themes and subcategories within themes from participant replies, this technique aims to "answer questions such as what, why, and how, and the common patterns in the data are searched for" [29].

Ethical Considerations

Informed consent was obtained from the physicians and nurse participating in the qualitative section of the study. All necessary approvals for this study were obtained from the Ethics Committees of Clalit Health Services and the University of Haifa (approval numbers 0031-16COM2 and 458/16, respectively).

Results

Telemedicine and Teletriage Growth Surge During the COVID-19 Pandemic

Telemedicine, or the use of digital and remote medical technologies to connect patients and caregivers, has become the hottest and most talked-about area of technology, thanks to the COVID-19 pandemic. The influence of the pandemic on the area of telemedicine worldwide is best summarized by the *New York Times* headline "10 Years of Change in One Week: Telemedicine on Fast Track" dated April 20, 2020.

COVID-19 plagued the world for most of 2020, posing a serious threat to public health. Although many health organizations were primarily focused on combating the immediate effects of COVID-19, maintaining basic and vital therapeutic services was equally important. Initial responses in many nations included clinic closures and the suspension of all noncritical medical services [30,31].

Telemedicine provides ongoing medical care, while maintaining strict social distancing. To reduce their exposure to others and still obtain medical care, patients at risk may benefit from staying at home. As a result, it is not surprising that health care

XSL•FO
RenderX

systems worldwide are turning to telemedicine, which has led to an exponential surge in its use as opposed to a previously slow uptake of the novel practice [32,33]. Thus, because of the COVID-19 pandemic, teletriage services have been implemented more frequently [34].

The benefits of teletriage during the COVID-19 pandemic have been described in recent studies; these studies show that this technique removes face-to-face contact, lowers the danger of exposure for medical personnel and other patients, and conserves scarce resources. Results suggest that more investigation is needed to ascertain how teletriage affects clinical outcomes, expenditures, and the use of follow-up care [35,36].

Although the COVID-19 pandemic has fueled the awareness and growth of technology and televisits, which are a convenience and infection control, the COVID-19 period has not made teletriage systems safer. It has made technology proliferate explosively.

Teletriage: First Point of Access to Care

Patients call advice lines for a reason. They want to know whether their symptoms are urgent. Clinical decision makers assess the symptoms, estimate the urgency, triage the symptoms, and advise when, where, and why the patient should be seen. Teletriage is designed specifically for this purpose—estimating symptom urgency and triage to ensure timely access to care. On the continuum of care, teletriage can now be acknowledged as the entry point to clinical care. It legitimately qualifies as "prehospital care" [8].

The primary function of teletriage is the assessment and management of symptoms by telephone, which also calls for expert judgment, clinical evaluation, and proactive information gathering from the patient [6,37].

According to researchers, nurses estimate and rule out symptom urgency to determine a disposition by using pattern recognition. "Telephonic medical diagnosis of patients' problems" is what telephone medicine, as practiced by doctors, is defined as [15,38].

Teletriage System Safety

The task in teletriage is to safely assess symptoms, estimate the urgency, and triage the symptoms presented remotely and then advise a disposition (time and place) for them to be further evaluated. The goal is to "make good decisions under conditions of uncertainty and urgency" to avoid the risk of delay in care, diagnosis, or treatment. Compared to in-person consultations, teletriage is a complex activity that entails certain inherent dangers because there is no visual contact and no nonverbal communication [39-41].

While performing teletriage, nurses must rely on audio signals rather than visual ones, although patients can speak about their symptoms using different terms. The ability of clinicians to communicate effectively is crucial, but there are also several other abilities that must be present, including the ability to recognize verbal cues, concentrate on obtaining a focused history, and understand the importance of having proper documentation [14,42,43].

Other characteristics of after-hours care that could pose risks include a high patient call volume, a variety of clinical conditions presented, the likelihood of urgent conditions being present, unknown patients, knowledge gaps regarding patients' medical histories, and the potential for information transfer discontinuity. Concerns have been raised because teletriage might compromise patient safety [44-48].

Regarding the reliability and safety of teletriage services, several recent studies have produced contradictory findings. Some studies were pessimistic, reporting that patient safety is frequently jeopardized by teletriage decisions [49]; service providers do not always forward the case to the on-call physician, when necessary [50]; and only a small number of diagnostic and therapeutic choices made during teletriage consultations offer the same level of health care as in-person conversations [51]. Inadequate visual cues that help doctors identify patients in acute condition were indicated as patient safety hazards in a study using teletriage [40].

However, more reassuring findings have been reported by other studies on the safety of teletriage systems. For instance, Blank et al [52] reviewed studies in which telephone counsel was contrasted with professional advice that is thought to be acceptable in that circumstance (ie, the "gold standard" of professional advice). The accuracy/appropriateness rate was 44%-98% in this review, with a median of 75%. Most decisions were appropriate according to a different study [14].

Concerning teletriage system effectiveness, the evidence also points to a variety of outcomes. According to certain studies, teletriage interventions, particularly for parents of small children and for older patients with chronic diseases, significantly reduce the number of emergency visits and readmissions [53,54]. Additionally, patients have stated that teletriage services have gained their trust and satisfaction. One study, however, found that a significant portion of patients who were directed to the ED using teletriage may have been treated elsewhere [55].

Based on a summary of several systematic reviews, when considered as a whole, the available research does not offer conclusive answers to queries concerning the standard of care delivered, the equity of access, costs, or outcomes in teletriage settings [18].

Growth alone in a new subspecialty will not guarantee safety. Developing a safe system is essential to any subspecialty, especially teletriage and telehealth. Defining the new subspecialty is one of the first challenges and sets the stage for transparency and, later, safety [14].

Even with the use of video and other technologies, remote symptom assessment is a uniquely risky task. Fraught with uncertainty, and many unknowns, teletriage is extremely time driven and time sensitive. A delay in care can be lethal if a required follow-up evaluation and treatment are not performed in a timely manner. In addition, teletriage is still in an underdeveloped state and lacks a reliable system. Finally, nurses and physicians perform this decision-making task under surprisingly difficult conditions [14,43-46].

Human factors in teletriage that challenge and possibly impair clinicians' decision-making process are detailed in Textbox 1.

XSL•FO
RenderX

Textbox 1. Telehealth risks (human factors).

- Inability to see patients (technology dependent)
- Ability to see but not to touch or gather patient vital signs (technology dependent)
- Extreme brevity of patient encounters (5-15 minutes)
- Incomplete or inaccurate information provided by patients
- Extensive sensory deprivation (endured by clinicians; technology dependent)
- Physical and cognitive demands imposed by high call volumes
- Potential for decision fatigue due to call volume and repetitive nature of the task [26,56]
- Clinicians often not knowing the patient, their education level, or their likelihood of compliance with advice
- A lack of structure (standardization of process and structure)
- Institutional pressures on clinicians to act as a gatekeeper rather than an access facilitator
- User-unfriendly electronic and paper guidelines

One way to avoid the risk of delays in care is to create a system. The Donabedian model [57-59] provides a framework for examining and evaluating health service quality. According to Donabedian [57-59], information about the quality of care can be drawn from 3 categories: structure, process, and outcomes.

Like other subspecialties, teletriage requires certain components to support safety. These components include standards (policies and procedures); sufficient numbers of qualified, experienced clinicians; specialized clinical training in medical decision-making; evidence-based, transparent, user-friendly guidelines; and electronic medical records (EMRs), audiotapes, or written documentation.

System Components: Evidence of a Duty to Care

Not surprisingly, in malpractice cases (when an error has occurred and a patient has been injured or has died), expert witnesses for the plaintiff always request tangible evidence of the system [9]:

- Guidelines used in the call (paper or electronic, eg, computerized decision support system [CDSS])
- Qualified experienced clinicians: résumés of nurses who managed the call, adequate numbers of clinicians
- Standards or policies and procedures, including job descriptions and qualifications
- Call center standards
- Actual call documentation: EMRs, paper form, or transcription of audiotaped calls
- Clinical teletriage training program materials

System Error

System error is thought to be the worst form of medical error [26]. Determining the effect of safety requires an examination of the problem of system error, defined as a failure of systems, processes, or conditions that are intended to prevent errors from occurring and that might lead people to make mistakes [60].

The Institute of Medicine (IOM) [60] has broadly defined system error as the “wrong match of plan” or the “failure to use any plan” to prevent error. For example, IOM research shows that the after-hours time, when no system in place, is especially risky in the United States [60]. In telehealth, complete systems (process or structure) are a first step toward reducing system error. Complete does not imply evidence-based or quality systems, however [8].

Malpractice in Teletriage

When a patient is harmed through unsafe telepractice, a malpractice case ensues. The plaintiff’s expert witnesses request evidence of care for that event: all documents that provide evidence of an adequate system, as described before. Institutions that can produce evidence of care are more able to demonstrate fulfillment of the duty of due care.

Physician teletriage malpractice may be related to the lack of a basic, complete teletriage system [16,49,61-65]. Nurse teletriage malpractice may be related to both the lack of a complete system or practicing in a complete system made up of faulty components [6,8,44,66-68].

What Are Meaningful Outcome Measures?

“We don’t look for patterns of our recurrent mistakes, or devise and refine potential solutions for them. But we could, and that is the ultimate point” [69].

We know what error and near misses look like. However, we have not yet clearly defined what constitutes safe practice and outcomes. Many researchers define telehealth safety variably, based on medical consensus on a study-by-study basis. Research continues to focus on nonessential elements of the process or structure (ie, communications, type of practitioner, patient compliance, and satisfaction).

The unfortunate outcomes described in malpractice [70-73] serve as fragments of the larger picture—system error, the essential and underrecognized problem.

Historically, medicine and nursing adhere to the key obligation “First, do no harm.” Nonmaleficence, which is derived from the maxim, is one of the principal precepts of bioethics—a fundamental principle worldwide.

Currently, professional organizations, such as the ANA [23] and the American Medical Association (AMA), typically set standards to guide medical decision-making, ethical practice, and patient safety. Formal systems—evidence-based structures

XSL•FO
RenderX

and methods and guidelines—support clinicians' safe practice and promote safe outcomes. Such system components are evolving slowly.

Safety Studies on US Teletriage

Research on the safety of teletriage systems, whether practiced by registered nurses (RNs) or MDs, is scarce [54].

Safety Research in the United States

Early studies examining the system structure and process provide a basis to inform research on system error. Although safety is often a topic of telehealth research, to the best of our knowledge, system error is still underresearched.

It is likely that the proprietary nature of telehealth technologies interferes with research on system safety. Telehealth trends make it difficult to achieve system transparency. The field urgently needs evidence-based CDSSs, EMRs, and other new technologies, such as features that provide feedback on outcomes to clinicians for the purpose of learning from their mistakes or successes.

In addition, CDSS, computerized decision-making system (CDMS) and EMR components, so fundamental to the clinical decision-making process, make it essential that these technologies be demonstrably and verifiably safe and effective. Questions remain about the safety of guideline technologies [74].

Early Research (1977-1990)

Early studies on teletriage focused on physician practice. Predictably, key demographic groups of frequent calls included infants and children, the elderly, and women. Topics also included categories of symptomatic calls and urgent situations: the sudden, rapid death of children, calls to the ED and poison centers, postpartum concerns, suicidal callers, and cases resulting in malpractice [9].

The first studies on remote telephone encounters often focused on problems that plagued physicians: strategies for reducing inappropriate after-hours calls, follow-up postdischarge calls, characteristics and perceptions of after-hours callers and high users ("frequent flyers"), call patterns, and dissatisfaction in pediatric practice. In general, US physicians were dissatisfied with the task of teletriage [9].

Research by Perrin and Goodman [75] marked the beginning of a change in how teletriage was practiced in the United States. The study compared the teletriage practices of pediatric nurse practitioners (PNPs) with those of pediatricians. Researchers found that PNPs are as safe and proficient as physicians, although PNPs take slightly more time to manage calls.

Research in 1990-2000

Research later focused on nurses' safety: communication, close calls, malpractice claims, access, chest pain, the influence of after-hours calls, and clinical and nonclinical decision makers. Later, the first teletriage training manual for nurses was published [9]. Lephrohon and Patel [14] showed how nurses practicing teletriage made decisions, describing pattern recognition and estimation of urgency as key decision-making strategies.

Research in 2000-2023

In the 2000s, rudimentary systems emerged [8]. Research highlighted the field's disorganization and lack of professional development [76]. Patient safety research was inconsistent and of variable quality, often commingling widely different clinical and nonclinical decision makers, intermingling terminology, and making unquestioned assumptions. Evidence-based studies were sparse.

A recently published systematic review [77] assessed the effectiveness of teletriage as one of these technologies during the COVID-19 pandemic. Studies investigating teletriage's effect on patient safety, clinical outcomes, and patient satisfaction were included. The authors concluded that teletriage interventions reduce unnecessary visits, improve clinical outcomes, reduce mortality and injuries, increase patient satisfaction, reduce health care provider workload, improve access to primary care consultation, and increase patient safety and satisfaction.

In Multimedia Appendix 3, we describe a developed teletriage center in the United States and include an interview with a qualified nurse working in this call center. Throughout the interview, she describes her personal feelings and reflections. Table 2 describes the required education, key system components, decision-making strategies, and goals of both Israeli physicians and US nurses.

XSL•FO
RenderX

Table 2. Decision maker comparison: Israeli MDs[a] and US RNs[b,c]

Decision maker	Minimum qualifications	System components	Decision-making strategies	Task objectives
Physician (autonomous, licensed clinician)	Doctorate level: 7 years of science-based clinical education and training + pediatrics specialty training for 4.5 years	Documentation: • Regulation: state medical board clinical training, guidelines	• Diagnosis • Clinical judgment • Critical thinking	• Make a medical diagnosis. • Identify and verify emergencies and urgencies.
Licensed nurse (autonomous, licensed clinician supported by a medically developed CDSS[d])	Associate of arts (AA)/bachelor of science (BS)/master of science (MS)/doctor of nursing practice (DNP): 2-7 years science-based clinical education and training	≥3 components • Guidelines: CDSS • Documentation: EMRs[e]/audiotaping, clinical teletriage training • Practice standards: American Academy of Ambulatory Care Nursing (AAACN) • Call center standards: URAC[f] • Regulation: Board of Registered Nursing	• Pattern recognition • Clinical judgment • Contextual information • Nursing process • Critical thinking	• Identify and verify emergencies and urgencies. • Estimate symptom urgency. • Rule out symptom urgency. • Interpret patient responses.

[a]MD: doctor of medicine.

[b]RN: registered nurse.

[c]Partially adapted from Wheeler [8], with permission.

[d]CDSS: computerized decision support system.

[e]EMR: electronic medical record.

[f]URAC: Utilization Review Accreditation Commission.

Telemedicine and Teletriage in Israel

In Israel, most of the health care and social assistance is public, including health care, welfare, child support, and old age and disability benefits. The national mandatory statutory health insurance system used in Israel is based on the Bismarck model. Both designated and ordinary taxes are used to pay for it. All citizens are required to join 1 of the 4 health plans (also known as mutualities or sick funds). The health plans provide both insurance for their members and a public basket of services, either through operating their own services or entering into contracts with service providers [78-80]. All 4 health plans are fully computerized, and all doctors and most other health care providers use EMRs that either are directly linked to the central medical record of the health plan through the internet or comprise its whole internal system. Between all community services, there is practically complete clinical data sharing. Highly developed decision support systems help with these.

Each health plan has highly advanced personal health records that allow members to access their own medical data online. These data entail prescription drug purchases and visits to the doctor, as well as imaging, laboratory, and other diagnostic test findings. Most of this is presently available online and via a smartphone in at least 2 of the health plans. Based on medical data and protocols created by the health plans, these plans currently provide proactive warnings and reminders for their members. The doctors at Maccabi, the second-largest health plan, can view their computerized medical information using a smartphone [78].

In Israel, physicians typically provide for all telehealth services, referred to as telemedicine. The physician practice of medicine or telemedicine is a range of remote high-tech remote encounters. The Ministry of Health (MOH) in Israel has regulations that apply to telemedicine services. Telemedicine standards were released in 2012 and have since been revised, as necessary, for different medical specialties. The MOH [79] provided an update in 2019 that details requirements for providing medical care remotely.

Although the worldwide pandemic has significantly accelerated what appears to be the next digital medical revolution, Israel has long recognized the enormous potential of telemedicine and has made it a national priority by allocating significant resources, establishing pertinent regulations, and promoting partnerships between health organizations, research institutions, start-up businesses, and independent researchers.

“Digital Health as an Engine of Growth” is a national priority program that Israel declared in March 2018. By using the information and communication technologies that are readily available to the entire Israeli population, the Israeli MOH [80] has stated that it is its mission to “bring about a leap in the health system that will enable it to become sustainable, advanced, innovative, renewed, and constantly improving.” In other words, the opportunity to further implement and expand a variety of telemedicine solutions is created by the worldwide acceleration of technology development and the digital revolution. The realization of the significance of digital health for the efficiency of the health care system and the requirement to offer strategic, systemic, and all-encompassing solutions for the foreseeable future are embodied in this national priority program [80].

Israel benefits greatly from a mix of human resources, a sizable number of businesses engaged in the development of digital medicine, and a sizable investment in research and development (R&D). It is a leader in communications and cyber innovation,

XSL•FO
RenderX

which is essential to the creation of cutting-edge digital medicine that will be used worldwide.

The conditions for the successful implementation of telemedicine in Israel are encouraging: the population has individual identification numbers, digital medical records are stored in sizable databases, all people have access to medical insurance, the standard of medicine is high, and communication technologies are of high quality and are widely available throughout the country [81].

In Israel, all health plans operate telemedicine services in one form or another. For administrative requirements with the clinic and the attending personal physician, they all permit online services. With each of them, the attending physician can also be reached via telephone or video call during clinic hours and sometimes even after hours.

Additionally, several of the health plans offer online pediatric and family services that primarily act as medical triage after working hours, throughout the evenings, nights, and weekends. The patients can use telephone or video calls and occasionally even submit images during the online consultations [82,83].

Some health plans have also begun using the TytoCare test device, which enables online physical assessment. During a digital visit, the equipment checks the patient's heart rate, respiration, temperature, ears, throat, and skin lesions using a variety of medical devices. A few Israeli hospitals have already begun to offer telemedicine consultations, particularly for presurgery evaluations, follow-up care, genetic and dietitian consultations, and even remote rehabilitation.

The quality of the telemedicine service provided and its safety are now the 2 most important factors to consider. Some telemedicine promotion initiatives during the pandemic seem to be predicated on the idea that a sizable part of outpatient visits may be effectively managed remotely, and patients can be prioritized for telemedicine services without endangering their safety or the standard of care [84].

An Israeli study [85] emphasized the growth in telemedicine usage during the first COVID-19 lockdown in Israel, as well as the anticipated partial fall in usage following the pandemic's end. As of May 2020, most Israeli pediatricians recommended that once the pandemic has passed, they return to in-person consultations and base their therapeutic judgments on frontal data rather than on data obtained through telemedicine contacts [85].

There are not many studies on the safety of telemedicine or teletriage services conducted in Israel. Haimi et al [84] examined the level of safety of a pediatric telemedicine service, paying particular attention to the accuracy of the diagnoses and the reasonability of judgments made by the online doctors. This service serves as a time-sensitive teletriage of spontaneous calls from parents about acute, worrisome symptoms of their children that require triage (symptom sorting). The study showed high levels of diagnosis accuracy (98.5%) and decision reasonableness (92%).

In addition to the literature review, using a qualitative study, we interviewed 15 physicians who had worked at the Clalit Pediatricians Online Service (a teletriage service) over the past 5 years [82-84,86]. Using a semistructured interview protocol form, we questioned the physicians about the difficulties and obstacles they face in the teletriage setting that may affect their capability in maintaining patient safety. In addition, they were asked about their perceptions of their capacity to uphold patient safety in this teletriage environment and, in particular, regarding elements that impacted their capacity to make reasonable decisions, determine the best course of action, and diagnose accurately, while upholding patient safety.

The physicians described several difficulties they face in the teletriage setting that may impact their ability to maintain patient safety [84]. The main factor was the difficulty to make a diagnosis from a distance due to the physician's inability to perform a physical examination in the telemedicine setting. Additional factors were treating unfamiliar patients, working alone, working under stressful conditions, having technological difficulties, and having a moral conflict between their desire to please and provide parents with good service on the one hand and the wish to maintain good medical practices on the other. While describing the challenges they face the teletriage setting, the physicians described various techniques and tools that they use to ensure patient safety.

Using a thematic analysis, we used the participants' replies to determine themes. These themes were compared with the original transcriptions to determine whether they accurately reflected the original data, guaranteeing a constant flow. The following themes were gleaned from the interviews with the 15 physicians:

- Use of intuition: Many physicians claimed to have used their intuition during the diagnostic process and frequently in relation to parents.

  You learn to rely on your intuition … whether you feel that the parents understand what you are saying, or that in this case, your instructions won't help.

  There is adversity, especially regarding certain decisions—I am sometimes hesitant about what to do, since I'm alone, especially at nights, and have to rely a lot on my intuition.

- Expertise: Most medical professionals believe that their clinical expertise in pediatrics in general and in telemedicine in particular aids in their diagnostic and decision-making processes. The more experience a medical professional has in telemedicine, the more confident they feel.

  During my first few days at work, I was afraid I would miss things or that there would be problems. After a while, however, I began to work with more confidence and less stress.

  There are some difficult aspects. At first, I felt insecure, but over time I gained experience (even the ability to diagnose better than the face-to-face doctor)! Like diagnosing a child with diabetic ketoacidosis …

- Using protocols: Many physicians said they use protocols and rules of thumb when making decisions. Most also use the protocols that are generated for special circumstances.

XSL•FO
RenderX

They believed this assisted them in maintaining patient safety. They were also conscious of potential biases in their thinking.

I use protocols. For example, head injuries among babies under the age of six months, or a high fever among babies younger than one month old. These make it easier to make a decision.

I use some rules of thumb. For instance, if a young boy is able to jump around, then he does not have appendicitis.

- Making shared decisions with parents: A few medical professionals reported talking to the parents of their patients about their opinions on the diagnostic process and potential treatment options.

I used to share my decision-making process with the parents. If there were several options, I would let the parents decide. In such a case, I depend on them.

I usually share, but I do not consult. I give my opinion and explain it, and only then do I wait for feedback.

- Using nonmedical factors: Most of the physicians agreed that they consider nonmedical considerations, in addition to medical factors when making decisions. Their opinions of the parents, particularly their level of comprehension, anxiety level, health literacy level, and the assurance that the parents will act appropriately if the child's illness worsens, are the most important considerations. The family's ability to access medical care was another crucial nonmedical element.

In addition to medical factors, the parents' tone of voice and level of stress may affect my decision, even if it seems to be a simple diagnosis … Language is also a factor. For example, new immigrants do not always understand me, and I am therefore more prone to sending them to the ED …

Aside from the medical condition, the patient's place of residency is also important. Living far from a medical care facility is a factor, and I will be more likely to consider an ED referral. In such cases, I also ask more questions about the availability of the doctor nearby.

You have to trust the parents' information and rely on them to follow the instructions correctly. If I feel that the chances of me being understood are poor (due to a lack of understanding or oversophistication on the part of the parents), I will refer them to the ED more easily.

- Additional techniques: The physicians schedule video conversations with the parents in cases of diagnostic doubt, ask them to send digital images, or schedule a follow-up call a few hours later.

If I needed additional information, I would arrange a video call or a follow-up call at a later time. Rarely would I consult with a senior physician.

Despite the difficulty making the decision, pictures and videos often compensate for the lack of a physical examination … In one case, I managed to correctly diagnose a child with intussusception!

Despite the difficulties and obstacles mentioned by online doctors [79], many of the physicians surveyed in this study reported having generally positive experiences with their telephone assessments and feeling confident in their ability to conduct thorough assessments and reach the right treatment decisions.

The key conclusions, with examples and comparisons between the 2 systems, are shown in Table 3.

XSL•FO
RenderX

Table 3. Key conclusions derived from the findings.

Key topics	Findings
Specialized clinical training for teletriage tasks	
RHP[a]	The RHP does not provide formal specialized teletriage training for nurses. However, it requires formal training for its electronic algorithms. Physicians present lectures on various specialties for the nurses.
Clalit Health Services	Teletriage training for pediatricians is not available. The authors believe training would aid pediatricians in making safer decisions during online consultations.
Conclusion for both systems	Judging from the interviews with nurses and physicians, it appears that both systems' clinical training is not adequate and formal training would be beneficial. Clinical training for any new subspecialty is an essential safety measure. Research has shown that clinical preparation has the potential to build confidence, improve performance, and reduce error, while improving morale [70-72,75,82,83].
Electronic algorithms and protocols	
RHP	With rare exception, the RHP requires nurses to follow and heavily rely on electronic algorithms in decision making. This raises the question of whether the RHP's electronic algorithms function more as a CDMS[b] than as a CDSS[c] [73]. The nurse interviewed (Ms Finley) stated that the overreliance on algorithms discourages nurses' critical thinking and dampens her initiative to perform a more thorough preliminary symptom assessment and to promote interpersonal interactions.
Clalit Health Services	The Clalit system provides several written protocols for certain clinical scenarios, and physicians are encouraged but not required to use them. In our qualitative interviews, many physicians said they used protocols and rules of thumb when making decisions.
Conclusion for both systems	For both nurses and physicians, guidelines are a key decision support tool. In addition, guideline quality (validity and reliability) requires evidence-based research—long overdue in this risk-prone field.
Documentation	
RHP	The RHP system provides 2 methods for call documentation, an audiotape recording and an electronic paper trail—a record of the patient-clinician encounter derived from a given guideline. However, the documentation output is limited to a patient's yes/no responses to the algorithmic questions. The result is an anonymized history with few details or context specific to a given patient [26]. Finley stated that physicians who later evaluate patients on-site do not have a good sense of why the patients were advised to be seen urgently. The RHP later developed a new policy allowing nurses to use a free-text area to document a brief symptom history using standard questions to elicit more specific details and context. Quality assurance is further bolstered by audiotaping all calls for follow-up review.
Clalit Health Services	The Clalit system requires physicians to document calls, completed in the child's medical file. As a result, the personal physician can view the online consultation during business hours. However, the language used in the documentation is completely up to the individual physician.
Conclusion for both systems	The RHP "paper trail" appears safer and more complete. However, the documented output appears to introduce confusion into on-site follow-up encounters. Clalit Health Services' lack of standardized language requirement may interfere with communication and continuity of care—a professional principle. Both systems are inadequate and increase miscommunication—one of the most common, recurrent error in this field.
Clinical call center standards (policies and procedures): clinicians' knowledge and experience	
RHP	According to Finley's interview, the RHP appears to have no job requirements or job descriptions and according to its policy may hire inexperienced nurse graduates. New nursing graduates are a poor match for the medical decision-making task, which according to many experts, requires a minimum 5-year bedside experience.
Clalit Health Services	The Clalit system hires only certified pediatricians, even though their level of experience as pediatricians in general and as online physicians may vary greatly.
Conclusion for both systems	Experience is critical in decision-making. Both groups could benefit from improved standards for required experience and job qualifications.
Clinical call center standards (policies and procedures): call length (teletriage meeting duration)	
RHP	Although it is the customary role of management to develop call center policies and procedures (standards), at the RHP, staff nurses have developed a minimal number of standards. One is a maximum call length, while another is a closing reminder to callers to call back if symptoms worsen or change.
Clalit Health Services	The Clalit system does not place any constraints on session length. However, since physicians are paid "per consultation," it may be an incentive to process calls quickly, although using the best medical decision.

XSL•FO
RenderX

Key topics	Findings
Clinical call center standards (policies and procedures): patient outcome feedback	
RHP	RHP nurses operate in a vacuum regarding patient outcomes (follow-up diagnosis). Outcomes provide feedback and are a measure of patient safety. Feedback about one's decisions is essential to improved practice and one of the strongest risk management measures available [87]. The rationale for not providing feedback to nurses is based on the Health Insurance Portability and Accountability Act (HIPAA). This federal law does not prevent US physicians' access to patient outcomes, however.
Clalit Health Services	Clalit physicians have complete access to the outcomes of their calls. Learning of their mistakes or successes may improve their practice and safety.
Conclusion for both systems	Ignorance about outcomes of one's decisions has never been shown to improve practice. Feedback mechanisms, known as planned error recovery, not only allow practitioners to learn the final diagnosis and thus improve their practice but also may improve guideline design and quality.

[a]RHP: Redwood Healthcare Plan.

[b]CDMS: computerized decision-making system.

[c]CDSS: computerized decision support system.

Discussion

Principal Findings

This narrative review and analysis presented a glimpse into current teletriage safety by analyzing 2 established and representative systems in 2 countries: Israel and the United States. We examined each system to learn how developed each is, perform a comparative analysis of both systems' safety, and explore the elements that might influence safe practice and patient outcomes.

In the initial stage, we carried out a thorough analysis of papers pertaining to patient safety in teletriage scenarios. Current research yields conflicting results regarding the dependability and security of teletriage systems. Although some critics claim that teletriage decisions frequently endanger patient safety [40,49-51], other research claims that using teletriage systems results in better safety outcomes [14,52].

We also analyzed a clinical call center of a large national US HMO based on the responses of a representative advice nurse to an interview (Multimedia Appendix 3), highlighting areas of risk that may contribute to system error [17]. We found that this representative system is still underdeveloped and lacks certain risk management elements. We based our conclusions on the interview, recent research, legal and risk management requirements related to the duty of due care, medical and nursing traditions, and existing subspecialty structures and processes.

In addition, we performed a qualitative study in which we interviewed 15 Israeli physicians working in a pediatric teletriage service in Israel, asking them about factors that affect their ability to maintain patient safety, while providing an accurate diagnosis, making appropriate decisions, and choosing the best course of action [83]. The physicians discussed the challenges they encounter in the telemedicine/teletriage context and the many strategies they use to arrive at the best diagnosis and course of care, protecting patient safety. These strategies include using their experience and intuition, using protocols generated for special clinical scenarios, making shared decisions with the patients (or their parents in the case of children), applying nonmedical criteria to aid in decision-making in situations where the medical data are ambiguous, and using more sophisticated tools (eg, video chats) when additional details are required. Many of the physicians surveyed in this study reported having generally positive experiences with their telephone assessments and feeling confident in their ability to conduct thorough assessments and make the best treatment decisions, despite the challenges and blockages described [82].

This study may be the first to examine and compare 2 official telehealth systems. For a combined 45-50 years, the 2 authors have performed triage in formal systems, taught, and provided consultation in the field of telehealth.

Teletriage, as stressed in this research, is the process of evaluating and prioritizing symptoms using telecommunication technologies. The main goal of teletriage is to assess and manage symptoms by telephone, which necessitates the use of professional judgment, clinical assessment, and proactive patient information gathering. The purpose of teletriage is to determine whether the needed on-site evaluation should take place and, if so, the venue and time. Teletriage involves clinical decision-making under remote and uncertain conditions. An overarching goal of teletriage safety is to avoid delays in care or diagnosis, which can cause patient harm.

Clinicians typically estimate the urgency of acute symptoms remotely and advise a disposition (triage level) for further medical diagnosis and treatment, as appropriate. The growth of teletriage services has accelerated due to the COVID-19 outbreak.

All types of health care delivery must consider safety, but with teletriage, this is both more crucial and challenging because acute symptoms may be time sensitive. Delay in care and diagnosis can result in harm to patients. Since there is no visible contact or nonverbal communication during teletriage, it is a more complicated activity than in-person consultations and it has certain inherent risks. The rapid pace of telehealth's growth creates urgency in identifying safe systems to guide developers and clinicians about needed improvement. Establishing a system is a key strategy to reduce the possibility of delay in care and diagnosis.

In the United States and internationally, one way to be cost-effective is to use the least paid person who can *safely* do

XSL•FO
RenderX

the job—an RN. Internationally, nurses have traditionally performed this task since the late 1980s. Early studies found that nurses are a safe substitute for physicians [14,73]. Thus, although physicians initially performed this task, they later delegated it to nurses.

Health care institutions historically provide standard features to support nurses and to enhance safety (subspecialty clinical training, standards, and documentation). In the case of teletriage, guidelines are typically written by physicians, similar to standing orders. These components provide a structure and process for this subspecialty and underpin safe practice.

An evolving subspecialty, even after 50 years, teletriage appears misunderstood and neglected. System error is thought to be a result of the absence or inadequacy of systems. In malpractice cases, expert witnesses for the patient or their family request evidence of the duty of due care. Typically, this evidence comprises documents: call documentation, guidelines used, clinical training materials, policies, and procedures (standards), including written job descriptions and qualifications.

Clearly, this analysis must acknowledge that contexts of the institutions described here differ in terms of respective health care systems and decision makers' clinical qualifications. The US health care system, and teletriage in particular, is plagued by disparate, competing forces: institutional cost containment, the need for professional standards, and diverging technological goals—the emphasis of speed over safety. This scenario requires better risk management.

Israel has universal health care, which appears to act differently. Physicians' depth and breadth of education and clinical training are superior to those of nurses. The US health care system compensates for this difference by providing more structure in the form of guidelines—typically developed collaboratively by physicians and software engineers. Physicians are not actual users of the guidelines that nurses are required to use.

Another variable is that of the populations served. Clalit pediatricians serve the needs of a diverse but still circumscribed pediatric population, whereas RHP nurses serve a broad, diverse population in terms of age range, symptom presentation, and diversity. This is a large order for nurses to manage and calls for a robust structure and process.

Finally, both RHP and Clalit systems share a common problem: incomplete systems of variable quality. The Clalit system's safety appears to rely on physician decision-making expertise, where standards, guidelines, and training are not that strong. The RHP may appear more complete. Safety may hinge on physician-developed electronic guidelines. Standards and training appear piecemeal or added as an afterthought. Without a meaningful, evidence-based structure and process in teletriage, quality (including safety) is at risk [18,58]. If establishing a system is a strategy to reduce possible error, then both systems could benefit from similar improvements.

Even if expert-level physicians require a less robust system, it appears that both physicians and nurses could benefit from specialized clinical training. In addition, consistent feedback regarding patient outcomes, known as planned error recovery—an essential error reduction strategy—promotes a method to self-check or to double-check another person's work [87].

Teletriage electronic algorithms must be evidence based. These guidelines are typically collaboratively developed by physicians and software developers. Nurses are required to use them, whereas physicians rarely use such tools.

Our narrative review and in-person interviews with physicians and a nurse about their experiences working in teletriage settings yielded several key findings, including the absence of specific formal training for the medical personnel working in teletriage; problematic protocols in particular clinical scenarios that, although not always available for all scenarios, are of low quality and do not allow for flexibility and agility, when needed; problematic documentation (mainly in nurse teletriage); inadequate experience and knowledge of the personnel who must make decisions in the face of uncertainty and urgency; limitations on the duration of calls or compensation based on the number of calls (which incentivizes personnel to conclude sessions promptly); and unsuitable feedback mechanisms that prevent personnel from understanding what transpired with patients and from learning from errors.

Drawing from our individual findings, the essential elements of teletriage are:

- Specialized clinical training for teletriage tasks
- Electronic algorithms and protocols
- Documentation
- Clinical call center standards: clinicians' knowledge and experience, call length (teletriage meeting duration), patient outcome feedback

Limitations

As with any narrative evaluation, selection bias cannot be completely ignored, even if this narrative analysis of the current literature was quite extensive and comprehensive and included a qualitative assessment of physicians and a nurse working in a teletriage setting.

Conclusion

Like other subspecialties, teletriage necessitates several elements to support safety, including qualified, experienced clinicians in sufficient numbers; specialized clinical training in medical decision-making; evidence-based, open, and approachable guidelines; and EMRs, audiotapes or written documentation, and standards (policies and procedures).

Fostering teletriage patient safety can be accomplished by taking the following general steps to improve MD and nurse practice in both Israel's and the United States' clinical call systems:

- Adequate training: Providers must receive adequate training to properly monitor and provide telehealth services. This includes knowledge of the systems being used, as well as familiarity with medical terminology and protocol.
- Regulation of telecommunication devices and systems: Providers must be aware of the regulations and requirements for the telecommunication devices and systems they use. This includes ensuring that the equipment is in good

XSL•FO
RenderX

working order and adheres to all safety and security regulations.
- Appropriate patient population: Telehealth services should only be used to treat patients who are stable and not at risk for an immediate life-threatening event. This will help ensure patient safety and avoid unwanted outcomes.
- Careful monitoring: In appropriate consultations, when needed, providers must carefully monitor patients and document any changes in their condition. This will help ensure that any changes or issues are addressed quickly and appropriately. Typically, nurses do not perform this task; in Israel, this is the role of the physician, for example, by using devices such as TytoCare.
- Quality assurance: Quality assurance protocols must be in place to ensure the accuracy and effectiveness of providers' services. This includes regularly reviewing documentation and providing feedback on any services deemed inadequate.
- Follow-up care: Providers must ensure that any patient receiving telehealth services receives follow-up care. This can include referrals to specialists or any other services needed to address any health concerns. Typically, nurses do not perform this task; in Israel, this is the role of the physician.
- Evidence-based studies of systems and safety: Misguided researchers unfamiliar with the triage task have produced confusing, misleading studies. Research that nibbles around the edges of the problem (patient or clinician satisfaction, clinician stress levels and attitudes, nonclinician practice) fails to address the core problem—system error. The telehealth industry requires long-overdue evidence-based outcome studies that meaningfully demonstrate the structures and processes that inform and strengthen safety.

Acknowledgments

We would like to express our gratitude to the physicians and the one nurse who took part in the interviews. We received no funding for this study.

Data Availability

All data analyzed during this study are included in this published paper and its supplementary information files. Additional data sets generated during this study are available from the corresponding author upon reasonable request.

Authors' Contributions

MH was involved in the conceptualization of the study, methodology, investigation, data curation, interviewing the physicians, analysis, and writing the paper. SQW was involved in data curation, methodology, formal analysis, interviewing the nurse, and writing the paper.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Definitions and terminology.
[DOCX File , 21 KB - humanfactors_v11i1e50676_app1.docx]

Multimedia Appendix 2
Paper selection criteria.
[DOCX File , 18 KB - humanfactors_v11i1e50676_app2.docx]

Multimedia Appendix 3
Case study from the United States.
[DOCX File , 30 KB - humanfactors_v11i1e50676_app3.docx]

References

1. Telemedicine market insights: the ultimate solution for quality care. Grand View Research. URL: https://tinyurl.com/4rhe6uyk [accessed 2024-03-14]
2. Grand View Research. URL: https://tinyurl.com/bp6k8n8r [accessed 2024-03-14]
3. Scott Kruse C, Karem P, Shifflett K, Vegi L, Ravi K, Brooks M. Evaluating barriers to adopting telemedicine worldwide: a systematic review. J Telemed Telecare 2018 Jan;24(1):4-12 [FREE Full text] [doi: 10.1177/1357633X16674087] [Medline: 29320966]
4. NEJM Catalyst. What is telehealth? New Engl J Med 2018 Feb 1;4(1):Online [FREE Full text] [doi: 10.1056/CAT.18.0268]
5. Global Market Insights. 2023 Jan. URL: https://tinyurl.com/2w36v6mn [accessed 2024-03-14]

XSL•FO
RenderX

6. Huibers L, Smits M, Renaud V, Giesen P, Wensing M. Safety of telephone triage in out-of-hours care: a systematic review. Scand J Prim Health Care 2011 Dec 29;29(4):198-209 [FREE Full text] [doi: 10.3109/02813432.2011.629150] [Medline: 22126218]
7. Boggan JC, Shoup JP, Whited JD, Van Voorhees E, Gordon AM, Rushton S, et al. Effectiveness of acute care remote triage systems: a systematic review. J Gen Intern Med 2020 Jul 02;35(7):2136-2145 [FREE Full text] [doi: 10.1007/s11606-019-05585-4] [Medline: 31898116]
8. Wheeler SQ, Greenberg ME, Mahlmeister L, Wolfe N. Safety of clinical and non-clinical decision makers in telephone triage: a narrative review. J Telemed Telecare 2015 Sep;21(6):305-322. [doi: 10.1177/1357633X15571650] [Medline: 25761468]
9. Wheeler SQ, Windt J. Telephone Triage: Theory, Practice and Protocol Development. San Anselmo, CA: TeleTriage Systems Publishers; 1993.
10. Wheeler SQ, Mahlmeister L. Telephone Triage Risk Management in Primary Care, Urgent Care and ED Settings. San Anselmo, CA: TeleTriage Systems Publishers; 2020.
11. Baren J, Rothrock S, Brennan J. In: Saunders WB, editor. Pediatric Emergency Medicine. Philadelphia, PA: Elsevier Health Sciences; 2008.
12. Taylor TB. COVID-19: rapid application of technology for emergency department tele-triage. American College of Emergency Physicians. 2020 Oct 6. URL: https://tinyurl.com/3bu2ncm5 [accessed 2024-03-14]
13. Huibers L, Keizer E, Giesen P, Grol R, Wensing M. Nurse telephone triage: good quality associated with appropriate decisions. Fam Pract 2012 Oct;29(5):547-552 [FREE Full text] [doi: 10.1093/fampra/cms005] [Medline: 22327415]
14. Leprohon J, Patel VL. Decision-making strategies for telephone triage in emergency medical services. Med Decis Mak 1995;15(3):240-253. [doi: 10.1177/0272989X9501500307] [Medline: 7564938]
15. Katz HP, Kaltsounis D, Halloran L, Mondor M. Patient safety and telephone medicine : some lessons from closed claim case review. J Gen Intern Med 2008 May;23(5):517-522 [FREE Full text] [doi: 10.1007/s11606-007-0491-y] [Medline: 18228110]
16. Nash DB. Insights into the use of digital tools for people with mental health disorders. Popul Health Manag 2021 Aug;24(S2):S53-S54. [doi: 10.1089/pop.2021.29004.dbn] [Medline: 34382865]
17. Kaminsky E, Röing M, Björkman A, Holmström IK. Telephone nursing in Sweden: a narrative literature review. Nurs Health Sci 2017 Sep 15;19(3):278-286. [doi: 10.1111/nhs.12349] [Medline: 28618087]
18. Lake R, Georgiou A, Li J, Li L, Byrne M, Robinson M, et al. The quality, safety and governance of telephone triage and advice services - an overview of evidence from systematic reviews. BMC Health Serv Res 2017 Aug 30;17(1):614 [FREE Full text] [doi: 10.1186/s12913-017-2564-x] [Medline: 28854916]
19. Wheeler SQ. ED telephone triage: lessons learned from unusual calls. J Emerg Nurs 1989;15(6):481-487. [Medline: 2687540]
20. Health contact center accreditation. Utilization Review Accreditation Commission. 2019. URL: https://www.urac.org/programs/health-call-center-accreditation [accessed 2024-03-14]
21. American Academy of Ambulatory Care Nursing. In: Anglea T, Murray C, Mastal M, Clelland S, editors. Scope and Standards of Practice for Professional Telehealth Nursing, 6th Edition. Pitman, NJ: American Academy of Ambulatory Care Nursing; 2018.
22. Framework for protecting patients when emergency care is out-of-network. American College of Emergency Physicians. 2019. URL: https://tinyurl.com/nhjxvdem [accessed 2024-03-14]
23. Principles of nursing documentation. American Nurses Association. 2010. URL: https://tinyurl.com/46mdkyrm [accessed 2024-03-14]
24. Emergency Nurses Association. Emergency Nurses Association position statement. Telephone advice. J Emerg Nurs 1991 Oct;17(5):52A. [Medline: 1921090]
25. Glossary of terms. NANDA International. URL: http://www.nanda.org/nanda-i-resources/glossary-of-terms/ [accessed 2024-03-14]
26. Wachter R. The Digital Doctor, Hope, Hype and Harm at the Dawn of Medicine's Computer Age. New York, NY: McGraw Hill; 2015.
27. Institute of Medicine (US) Committee on Standards for Developing Trustworthy Clinical Practice Guidelines. In: Graham R, Mancher M, Wolman DM, Greenfield S, Steinberg E, editors. Clinical Practice Guidelines We Can Trust. Washington, DC: National Academies Press; 2011.
28. O'Brien BC, Harris IB, Beckman TJ, Reed DA, Cook DA. Standards for reporting qualitative research: a synthesis of recommendations. Acad Med 2014 Sep;89(9):1245-1251 [FREE Full text] [doi: 10.1097/ACM.0000000000000388] [Medline: 24979285]
29. Cho JY, Lee E. Reducing confusion about grounded theory and qualitative content analysis: similarities and differences. Qual Rep 2014;19(32):1-20 [FREE Full text] [doi: 10.46743/2160-3715/2014.1028]
30. Huang C, Wang Y, Li X, Ren L, Zhao J, Hu Y, et al. Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China. Lancet 2020 Feb 15;395(10223):497-506 [FREE Full text] [doi: 10.1016/S0140-6736(20)30183-5] [Medline: 31986264]

XSL•FO
RenderX

31. Ting DSW, Carin L, Dzau V, Wong TY. Digital technology and COVID-19. Nat Med 2020 Apr 27;26(4):459-461 [FREE Full text] [doi: 10.1038/s41591-020-0824-5] [Medline: 32284618]
32. Elkbuli A, Ehrlich H, McKenney M. The effective use of telemedicine to save lives and maintain structure in a healthcare system: current response to COVID-19. Am J Emerg Med 2021 Jun;44:468-469 [FREE Full text] [doi: 10.1016/j.ajem.2020.04.003] [Medline: 32303410]
33. Bashshur R, Doarn C, Frenk J, Kvedar J, Woolliscroft J. Telemedicine and the COVID-19 pandemic, lessons for the future. Telemed J E Health 2020 May;26(5):571-573 [FREE Full text] [doi: 10.1089/tmj.2020.29040.rb] [Medline: 32275485]
34. Smith AC, Thomas E, Snoswell CL, Haydon H, Mehrotra A, Clemensen J, et al. Telehealth for global emergencies: implications for coronavirus disease 2019 (COVID-19). J Telemed Telecare 2020 Mar 20;26(5):309-313. [doi: 10.1177/1357633x20916567]
35. Christianson J, Christianson E. White paper: using telehealth in the emergency department to minimize risk to health care providers and conserve resources during the COVID-19 response. InDoc, PA. 2020. URL: ttps://tinyurl.com/yc3h8nbe [accessed 2024-03-14]
36. Desborough J, Hall Dykgraaf S, de Toca L, Davis S, Roberts L, Kelaher C, et al. Australia's national COVID-19 primary care response. Med J Aust 2020 Aug 04;213(3):104-106.e1 [FREE Full text] [doi: 10.5694/mja2.50693] [Medline: 32623740]
37. Wheeler SQ. Telephone Triage Protocols for Adult Populations. New York, NY: McGraw Hill Publishers; 2009.
38. Reisman AB, Brown KE. Preventing communication errors in telephone medicine. J Gen Intern Med 2005 Oct;20(10):959-963 [FREE Full text] [doi: 10.1111/j.1525-1497.2005.0199.x] [Medline: 16191150]
39. Patel A, Dale J, Crouch R. Satisfaction with telephone advice from an accident and emergency department: identifying areas for service improvement. Qual Health Care 1997 Sep 01;6(3):140-145 [FREE Full text] [doi: 10.1136/qshc.6.3.140] [Medline: 10173771]
40. Salk ED, Schriger DL, Hubbell KA, Schwartz BL. Effect of visual cues, vital signs, and protocols on triage: a prospective randomized crossover trial. Ann Emerg Med 1998 Dec;32(6):655-664. [doi: 10.1016/s0196-0644(98)70063-0] [Medline: 9832660]
41. Smith S. Helping parents cope with crying babies: decision-making and interaction at NHS Direct. J Adv Nurs 2010 Feb 20;66(2):381-391. [doi: 10.1111/j.1365-2648.2009.05195.x] [Medline: 20423421]
42. Holmström I. Decision aid software programs in telenursing: not used as intended? Experiences of Swedish telenurses. Nurs Health Sci 2007 Mar 13;9(1):23-28. [doi: 10.1111/j.1442-2018.2007.00299.x] [Medline: 17300541]
43. Car J, Freeman GK, Partridge MR, Sheikh A. Improving quality and safety of telephone based delivery of care: teaching telephone consultation skills. Qual Saf Health Care 2004 Feb 01;13(1):2-3 [FREE Full text] [doi: 10.1136/qshc.2003.009241] [Medline: 14757786]
44. Giesen P, Ferwerda R, Tijssen R, Mokkink H, Drijver R, van den Bosch W, et al. Safety of telephone triage in general practitioner cooperatives: do triage nurses correctly estimate urgency? Qual Saf Health Care 2007 Jun 01;16(3):181-184 [FREE Full text] [doi: 10.1136/qshc.2006.018846] [Medline: 17545343]
45. Smits M, Huibers L, Kerssemeijer B, de Feijter E, Wensing M, Giesen P. Patient safety in out-of-hours primary care: a review of patient records. BMC Health Serv Res 2010 Dec 10;10(1):335 [FREE Full text] [doi: 10.1186/1472-6963-10-335] [Medline: 21143949]
46. Derkx HP, Rethans JE, Muijtjens AM, Maiburg BH, Winkens R, van Rooij HG, et al. Quality of clinical aspects of call handling at Dutch out of hours centres: cross sectional national study. BMJ 2008 Sep 12;337(sep12 1):a1264-a1264 [FREE Full text] [doi: 10.1136/bmj.a1264] [Medline: 18790814]
47. Fourny M, Lucas A, Belle L, Debaty G, Casez P, Bouvaist H, et al. Inappropriate dispatcher decision for emergency medical service users with acute myocardial infarction. Am J Emerg Med 2011 Jan;29(1):37-42. [doi: 10.1016/j.ajem.2009.07.008] [Medline: 20825772]
48. Moriarty H, McLeod D, Dowell A. Mystery shopping in health service evaluation. Br J Gen Pract 2003 Dec;53(497):942-946 [FREE Full text] [Medline: 14960218]
49. Killip S, Ireson CL, Love MM, Fleming ST, Katirai W, Sandford K. Patient safety in after-hours telephone medicine. Fam Med 2007 Jun;39(6):404-409 [FREE Full text] [Medline: 17549649]
50. Hildebrandt DE, Westfall JM, Smith PC. After-hours telephone triage affects patient safety. J Fam Pract 2003 Mar;52(3):222-227. [Medline: 12620177]
51. Hersh W, Helfand M, Wallace J, Kraemer D, Patterson P, Shapiro S, et al. A systematic review of the efficacy of telemedicine for making diagnostic and management decisions. J Telemed Telecare 2002;8(4):197-209. [doi: 10.1258/135763302320272167] [Medline: 12217102]
52. Blank L, Coster J, O'Cathain A, Knowles E, Tosh J, Turner J, et al. The appropriateness of, and compliance with, telephone triage decisions: a systematic review and narrative synthesis. J Adv Nurs 2012 Dec 07;68(12):2610-2621. [doi: 10.1111/j.1365-2648.2012.06052.x] [Medline: 22676805]
53. Nataraj J, Stempien J, Netherton S, Wahba MY, Oyedokun TO. Emergency department referrals from a provincial medical call centre: is it more than just 1-800-go-to-emerg? Can J Emerg Med 2020 Mar 24;22(2):241-244. [doi: 10.1017/cem.2019.420] [Medline: 31645231]

XSL•FO
RenderX

54. Sax DR, Vinson DR, Yamin CK, Huang J, Falck TM, Bhargava R, et al. Tele-triage outcomes for patients with chest pain: comparing physicians and registered nurses. Health Aff (Millwood) 2018 Dec;37(12):1997-2004. [doi: 10.1377/hlthaff.2018.05079] [Medline: 30633669]
55. Turbitt E, Freed GL. Use of a telenursing triage service by Victorian parents attending the emergency department for their child's lower urgency condition. Emerg Med Australas 2015 Dec 03;27(6):558-562. [doi: 10.1111/1742-6723.12477] [Medline: 26337783]
56. Baumeister RF, Tierney J. Willpower: Rediscovering the Greatest Human Strength. London, UK: Penguin Press; 2011.
57. Donabedian A. Evaluating the quality of medical care. Milbank Memorial Fund Q 1966 Jul;44(3):166. [doi: 10.2307/3348969]
58. Donabedian A. Quality assurance. Structure, process and outcome. Nurs Stand 1992;7(11 Suppl QA):4-5. [Medline: 1489693]
59. Berwick D, Fox DM. "Evaluating the quality of medical care": Donabedian's classic article 50 years later. Milbank Q 2016 Jun;94(2):237-241 [FREE Full text] [doi: 10.1111/1468-0009.12189] [Medline: 27265554]
60. Institute of Medicine (US) Committee on Quality of Health Care in America. In: Kohn LT, Corrigan JM, Donaldson MS, editors. To Err is Human: Building a Safer Health System. Washington, DC: National Academies Press; 2000.
61. Scarfone RJ, Luberti AA, Mistry RD. Outcomes of children referred to an emergency department by an after-hours call center. Pediatr Emerg Care 2004 Jun;20(6):367-372. [doi: 10.1097/01.pec.0000133610.42699.41] [Medline: 15179144]
62. Hildebrandt DE, Westfall JM, Fernald DH, Pace WD. Harm resulting from inappropriate telephone triage in primary care. J Am Board Fam Med 2006 Sep 01;19(5):437-442 [FREE Full text] [doi: 10.3122/jabfm.19.5.437] [Medline: 16951292]
63. Wetterneck TB, Karsh BT. Human factors applications to understanding and using close calls to improve health care. In: Wu A, editor. The Value of Close Calls in Improving Patient Safety: Learning How to Avoid and Mitigate Patient Harm. Oak Brook, IL: The Joint Commission Resources; 2011:39-53.
64. Makary MA, Daniel M. Medical error—the third leading cause of death in the US. BMJ 2016 May 03;353:i2139 [FREE Full text] [doi: 10.1136/bmj.i2139] [Medline: 27143499]
65. Walker J. EHR: a MedMal minefield for practitioners? JLNC 2018;29(2):9-13 [FREE Full text] [doi: 10.30710/JLNC.29.2.2018.9]
66. Hirsh D, Simon H, Massey R, Thornton L, Simon J. The host hospital 24-hour underreferral rate: an automated measure of call-center safety. Pediatrics 2007 Jun;119(6):1139-1144. [doi: 10.1542/peds.2006-1986] [Medline: 17545381]
67. Mahlmeister L. Human factors and errors in perinatal care. J Perinat Neonat Nurs 2015;24(1):12-21. [doi: 10.1097/jpn.0b013e3181cdb6e0]
68. Mahlmeister L, Van Mullem C. The process of triage in perinatal settings: clinical and legal issues. J Perinat Neonat Nurs 2000 Mar;13(4):13-30. [doi: 10.1097/00005237-200003000-00003] [Medline: 11075083]
69. Gawande A. The Checklist Manifesto: How to Get Things Right. New York, NY: Picador; 2010.
70. Ernesäter A, Engström M, Holmström I, Winblad U. Incident reporting in nurse-led national telephone triage in Sweden: the reported errors reveal a pattern that needs to be broken. J Telemed Telecare 2010 May 10;16(5):243-247. [doi: 10.1258/jtt.2009.090813] [Medline: 20457800]
71. Ernesäter A, Winblad U, Engström M, Holmström IK. Malpractice claims regarding calls to Swedish telephone advice nursing: what went wrong and why? J Telemed Telecare 2012 Aug 24;18(7):379-383. [doi: 10.1258/jtt.2012.120416] [Medline: 22923361]
72. Budryk Z. Malpractice claims against nurses on the rise. FIERCE Healthcare. 2016 Feb 26. URL: https://tinyurl.com/yn8t8764 [accessed 2024-03-14]
73. Oken RL. What have we learned from pediatric malpractice cases? AAP News. 2018. URL: https://publications.aap.org/aapnews/news/14845 [accessed 2024-03-14]
74. Greatbatch D, Hanlon G, Goode J, O'Cathain A, Strangleman T, Luff D. Telephone triage, expert systems and clinical expertise. Sociol Health Illn 2005 Sep;27(6):802-830 [FREE Full text] [doi: 10.1111/j.1467-9566.2005.00475.x] [Medline: 16283900]
75. Perrin EC, Goodman HC. Telephone management of acute pediatric illnesses. N Engl J Med 1978 Jan 19;298(3):130-135. [doi: 10.1056/nejm197801192980304]
76. Carrasqueiro S, Oliveira M, Encarnação P. Evaluation of telephone triage and advice services: a systematic review on methods, metrics and results. Stud Health Technol Inform 2011;169:407-411. [Medline: 21893782]
77. Farzandipour M, Nabovati E, Sharif R. The effectiveness of tele-triage during the COVID-19 pandemic: a systematic review and narrative synthesis. J Telemed Telecare 2023 Jan 23:1357633X221150278 [FREE Full text] [doi: 10.1177/1357633X221150278] [Medline: 36683438]
78. International health care system profiles. The Commonwealth Fund. 2019. URL: https://tinyurl.com/ycym3fcb [accessed 2024-03-14]
79. Discussed: criteria for operating a remote health service. Ministry of Health Israel. URL: https://www.health.gov.il/hozer/mk06_2019.pdf [accessed 2024-03-14]
80. National plan for digital health. Ministry of Health Israel. URL: https://www.health.gov.il/About/projects/DigitalHealth/Pages/default.aspx [accessed 2024-03-14]

XSL•FO
RenderX

81. Gamus A, Chodick G. Telemedicine after COVID-19: the Israeli perspective. Isr Med Assoc J 2020 Aug;22(8):467-469 [FREE Full text] [Medline: 33236577]
82. Haimi M, Brammli-Greenberg S, Waisman Y, Baron-Epel O. Physicians' experiences, attitudes and challenges in a pediatric telemedicine service. Pediatr Res 2018 Nov;84(5):650-656. [doi: 10.1038/s41390-018-0117-6] [Medline: 30120402]
83. Haimi M, Brammli-Greenberg S, Waisman Y, Stein N, Baron-Epel O. The role of non-medical factors in physicians' decision-making process in a pediatric telemedicine service. Health Informat J 2020 Jun;26(2):1152-1176 [FREE Full text] [doi: 10.1177/1460458219870660] [Medline: 31566448]
84. Haimi M, Brammli-Greenberg S, Baron-Epel O, Waisman Y. Assessing patient safety in a pediatric telemedicine setting: a multi-methods study. BMC Med Inform Decis Mak 2020 Apr 03;20(1):63 [FREE Full text] [doi: 10.1186/s12911-020-1074-7] [Medline: 32245469]
85. Grossman Z, Chodick G, Reingold SM, Chapnick G, Ashkenazi S. The future of telemedicine visits after COVID-19: perceptions of primary care pediatricians. Isr J Health Policy Res 2020 Oct 20;9(1):53 [FREE Full text] [doi: 10.1186/s13584-020-00414-0] [Medline: 33081834]
86. Nevet A, Bitton Y, Wolf L, Waisman Y. [Telemedicine: a novel service in pediatric emergency care]. Harefuah 2016 Jul;155(7):410-413. [Medline: 28514123]
87. Human factors analysis in patient safety systems. The Joint Commission. 2015 Apr. URL: https://tinyurl.com/38zjxcts [accessed 2024-03-14]

Abbreviations

ACEP: American College of Emergency Physicians
ANA: American Nurses Association
AR: appropriate referral
CDMS: computerized decision-making system
CDSS: computerized decision support system
ED: emergency department
EMR: electronic medical record
HMO: health maintenance organization
MD: doctor of medicine
MOH: Ministry of Health
OR: overreferral
PNP: pediatric nurse practitioner
RHP: Redwood Healthcare Plan
RN: registered nurse
UR: underreferral
URAC: Utilization Review Accreditation Commission

Edited by A Kushniruk; submitted 09.07.23; peer-reviewed by G Chodick, A Benis, T Aslanidis; comments to author 07.11.23; revised version received 25.11.23; accepted 28.02.24; published 25.03.24.

Please cite as:
Haimi M, Wheeler SQ
Safety in Teletriage by Nurses and Physicians in the United States and Israel: Narrative Review and Qualitative Study
JMIR Hum Factors 2024;11:e50676
URL: https://humanfactors.jmir.org/2024/1/e50676
doi:10.2196/50676
PMID:38526526

©Motti Haimi, Sheila Quilter Wheeler. Originally published in JMIR Human Factors (https://humanfactors.jmir.org), 25.03.2024. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on https://humanfactors.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Evaluating the Usability of an mHealth App for Empowering Cancer Survivors With Disabilities: Heuristic Evaluation and Usability Testing

Rachel F Adler[1,2], PhD; Kevin Baez[1], BS; Paulina Morales[1], BS; Jocelyn Sotelo[1], BS; David Victorson[3], PhD; Susan Magasi[4], PhD

[1]Department of Computer Science, Northeastern Illinois University, Chicago, IL, United States
[2]School of Information Sciences, University of Illinois Urbana-Champaign, Champaign, IL, United States
[3]Department of Medical Social Sciences, Northwestern University, Evanston, IL, United States
[4]Department of Occupational Therapy, University of Illinois Chicago, Chicago, IL, United States

Corresponding Author:
Rachel F Adler, PhD
School of Information Sciences
University of Illinois Urbana-Champaign
501 E. Daniel St.
Champaign, IL, 61820
United States
Phone: 1 217 244 4192
Email: radler@illinois.edu

Abstract

Background: More than 18 million cancer survivors are living in the United States. The effects of cancer and its treatments can have cognitive, psychological, physical, and social consequences that many survivors find incredibly disabling. Posttreatment support is often unavailable or underused, especially for survivors living with disabilities. This leaves them to deal with new obstacles and struggles on their own, oftentimes feeling lost during this transition. Mobile health (mHealth) interventions have been shown to effectively aid cancer survivors in dealing with many of the aftereffects of cancer and its treatments; these interventions hold immense potential for survivors living with disabilities. We developed a prototype for WeCanManage, an mHealth-delivered self-management intervention to empower cancer survivors living with disabilities through problem-solving, mindfulness, and self-advocacy training.

Objective: Our study conducted a heuristic evaluation of the WeCanManage high-fidelity prototype and assessed its usability among cancer survivors with known disabilities.

Methods: We evaluated the prototype using Nielsen's 10 principles of heuristic evaluation with 22 human-computer interaction university students. On the basis of the heuristic evaluation findings, we modified the prototype and conducted usability testing on 10 cancer survivors with a variety of known disabilities, examining effectiveness, efficiency, usability, and satisfaction, including a completion of the modified System Usability Scale (SUS).

Results: The findings from the heuristic evaluation were mostly favorable, highlighting the need for a help guide, addressing accessibility concerns, and enhancing the navigation experience. After usability testing, the average SUS score was 81, indicating a good-excellent design. The participants in the usability testing sample expressed positive reactions toward the app's design, educational content and videos, and the available means of connecting with others. They identified areas for improvement, such as improving accessibility, simplifying navigation within the community forums, and providing a more convenient method to access the help guide.

Conclusions: Overall, usability testing showed positive results for the design of WeCanManage. The course content and features helped participants feel heard, understood, and less alone.

(JMIR Hum Factors 2024;11:e51522) doi:10.2196/51522

XSL•FO
RenderX

KEYWORDS

mobile health; mHealth; apps; usability; cancer survivors; accessibility; disabilities; cancer; oncology; heuristics; empowerment; advocacy; mindfulness; problem-solving

Introduction

Background

There are an estimated 18.1 million cancer survivors in the United States, and the number is projected to increase to 22.5 million by 2032 [1]. Approximately 40% of cancer survivors experience long-term physical, cognitive, psychological, and social consequences of cancer and its treatment, which can lead to significant disability [2]. These effects can include physical challenges, including but not limited to pain, fatigue, decreased functional mobility, limb loss, lymphedema, speech and swallowing difficulties, emotional challenges (as cancer survivors may experience anxiety or depression), and cognitive challenges (such as "chemo brain") [3-5]. These aftereffects can lead to activity limitations and participation restrictions, which according to contemporary frameworks and legal definitions may be considered as disabilities [6,7]. Yet, even with significant functional impairments, not all cancer survivors self-identify as disabled [8,9]. Regardless of the terminology used, the aftereffects of cancer and their related functional impacts can have a significant negative impact on well-being and health-related quality of life [10]. Survivorship plans and rehabilitation programs, which play a crucial role in restoring survivors' physical and emotional well-being, are frequently underused by cancer survivors [11]. This can be due to obstacles like time, financial constraints, and transportation issues [12], which hinder their accessibility. Mobile health (mHealth) apps can help make rehabilitation services accessible and put them in the hands of those who need them.

mHealth Apps

Mobile technologies—smartphones, tablets, and smartwatches—are increasingly ubiquitous in today's society and can be used almost anywhere [13]. The Pew Research Center reports that 85% of American adults own smartphones, and the ownership is relatively consistent across genders; racial groups; and urban, suburban, and rural users [14]. This leads to an increase in the development of mHealth apps. The COVID-19 pandemic has led to mHealth strategies becoming even more important in cancer care. According to the recommendations of Curigliano et al [15], patients with cancer should be offered mHealth strategies to support symptom management and adoption of healthy behaviors. The number of mHealth apps has increased throughout the years, with around 325,000 apps available in 2017 [16]. Charbonneau et al [17] identified 123 mHealth apps for cancer survivors available in the 2 most important marketplaces (ie, Apple iTunes and Google Play). Typical areas of usage in cancer are disease management support (eg, symptom monitoring, management of side effects, medication reminder and dosing, and access to health information), support of healthy behavior (eg, healthy diet and increased physical activity), or the connection with other patients (eg, social support through peers) [18-20].

Evaluating the Usability of mHealth Apps

It is important to gather qualitative and quantitative data on mHealth apps to determine how satisfied users would be with the product at hand. According to one scoping review, of 133 different eHealth articles that conducted usability testing, 105 used questionnaires, 57 used task completion, 45 used "think aloud," 37 conducted interviews, 18 performed heuristic evaluation, and 13 used focus groups [21]. The System Usability Scale (SUS) was the most frequently used questionnaire with a total of 44 studies. A combination of methods was used in 88 of the studies. Further, cancer was tied as the second most frequently evaluated health condition (n=10), with only mental health being evaluated more often (n=12).

Usability testing is a common effective method for evaluating the usability of mHealth apps. Studies have shown that usability testing is an effective method for examining mHealth apps for diabetes [22,23], depression [22,24], and youth at risk for developing psychosis [25], as well as managing pain [26], heart failure [27], and cancer symptoms [28]. Common questionnaires often included variations on the Mobile Application Rating Scale [25,27] or the SUS [22,24,26]. Additional techniques often employed in usability testing include measuring time per task [26] and using think aloud techniques [29]. In addition to evaluating fully implemented mobile apps, studies have conducted usability testing on prototypes of mHealth apps for supporting mental health [30], chronic kidney disease [29], fall risk detection system for older users [31], HIV [32], and cancer survivors [33-35]. Many studies have conducted heuristic evaluation before usability testing on an mHealth prototype to fix usability issues before bringing it to users [28,29,32,33]. While Nielsen's 10-point usability heuristics [36] are geared toward computer-based applications, most of these are also applicable in mobile app design. The SUS questionnaire was also commonly used in usability testing studies for examining mHealth prototypes [29,31,37].

WeCanManage App

We designed a high-fidelity prototype for WeCanManage, an evidence-informed mHealth self-management intervention, aimed at empowering individuals with tools to effectively manage cancer as a chronic condition. Users are asked to log into the app daily for 5-10 minutes to complete mobile microlearning modules of self-management content. The intervention content is based on extensive literature review and formative interviews with cancer survivors with known disabilities (n=30) and supportive cancer care professionals including social workers, psychologists, occupational and physical therapists, and a physiatrist specializing in cancer rehabilitation (n=5) [9]. A team of survivor scientists, people with lived experiences of cancer and disability, further informed intervention content and focus. Intervention content is presented sequentially as information is scaffolded on itself to promote depth of learning, retention, and application. The content is divided into 4 broad sections: WeCanRelate (fosters a sense of

XSL•FO
RenderX

validating and normalizing the survivorship experience), WeCanAdapt (teaches goal direction self-management strategies), WeCanBe (emphasizes mindfulness-based practices), and WeCanSpeakUp (addresses self-advocacy and disability rights). In addition to the instructional content, WeCanManage provides users with 3 circles of support, including one-on-one connections with other users (Connect to Peers [C2P]), community forums (to discuss intervention content and shared experiences with the entire user community), and a library with evidence-informed educational content [38]. We conducted a thorough evaluation of the usability of the high-fidelity prototype for cancer survivors with disabilities, employing both heuristic evaluation and usability testing to assess its effectiveness in addressing the unique needs and challenges of this user group.

Methods

WeCanManage High-Fidelity Prototype

The high-fidelity prototype was created on Marvel [39], a web-based collaborative design platform that provides tools for creating wireframes, designs, and prototypes of interactive applications. We aimed to design WeCanManage specifically for smartphone usage. The prototype of WeCanManage allows users to navigate between the Home, Journey (Courses), C2P, Community (Community Forum), and Library (see Figure 1).

The Course section provides cancer survivors with an educational intervention that works with them on dealing with the long-term effects of their newly acquired disabilities through problem-solving, mindfulness, and self-advocacy. The content is designed to be a 4-week program where the user unlocks a series of microlessons divided into 4 modules (WeCanRelate, WeCanAdapt, WeCanBreathe, and WeCanSpeakUp), which educate users with different methods to deal with the effects of postcancer treatment in their daily life. To prioritize user control and accessibility, the course content is conveyed through mobile microlearning modules, presented in different formats such as readable text, clickable text-based cards, and audio (Figure 2).

At the end of many of the daily sessions, there are interactive engagement activities, such as reflections that feed into the Community Forum and knowledge checks (see Figure 3). The engagement activities are designed to support consolidation of knowledge and application of course content to the user's lived experiences.

The Community and C2P sections offer users a chance to engage with others, fostering networking opportunities and creating a support system with individuals undergoing similar experiences. C2P facilitates connections with others, allowing users to filter by categories like cancer type and disability, while Community features discussion forums for each of the 4 course sections and an open discussion forum. Lastly, the Library section contains additional evidence-informed resources such as articles and factsheets. The various sections of the prototypes were initially created as a low-fidelity prototype through an iterative co-design approach involving both the design teams and cancer survivors, who served as representatives of our targeted audience [40].

Because of its prototype nature, users could navigate all links, but functionalities such as real-time chat with other users and composing reflections or community posts were not operational. To overcome this, we incorporated simulated features in the prototype, triggering them automatically on user interaction. After creating the high-fidelity prototype, we evaluated it through 2 distinct methods: heuristic evaluation and usability testing.

Figure 1. Screenshots of the WeCanManage prototype: (A) Home, (B) Journey, (C) Connect to Peers (C2P), (D) Community, and (E) Library.

XSL•FO
RenderX

Figure 2. Screenshots before heuristic evaluation: (A) card view and (B) learning format after clicking on the Formats icon.

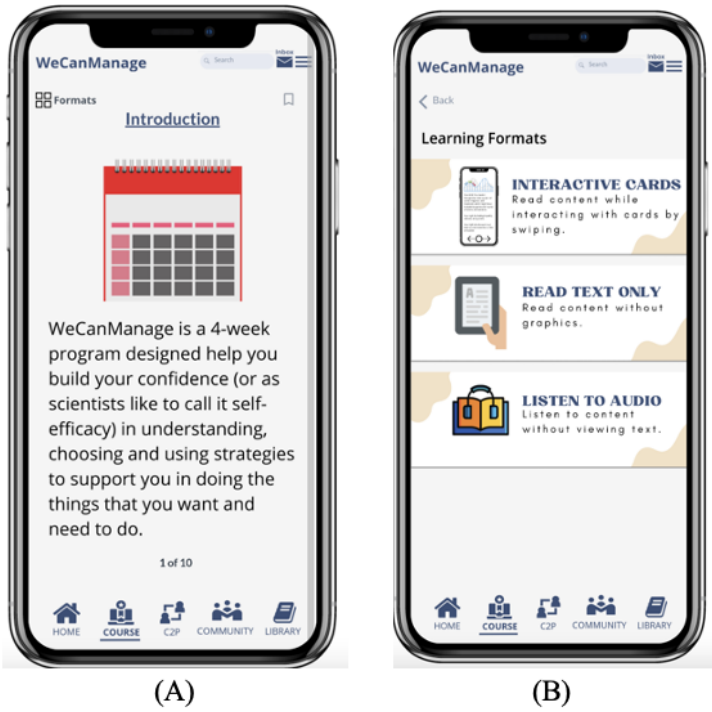
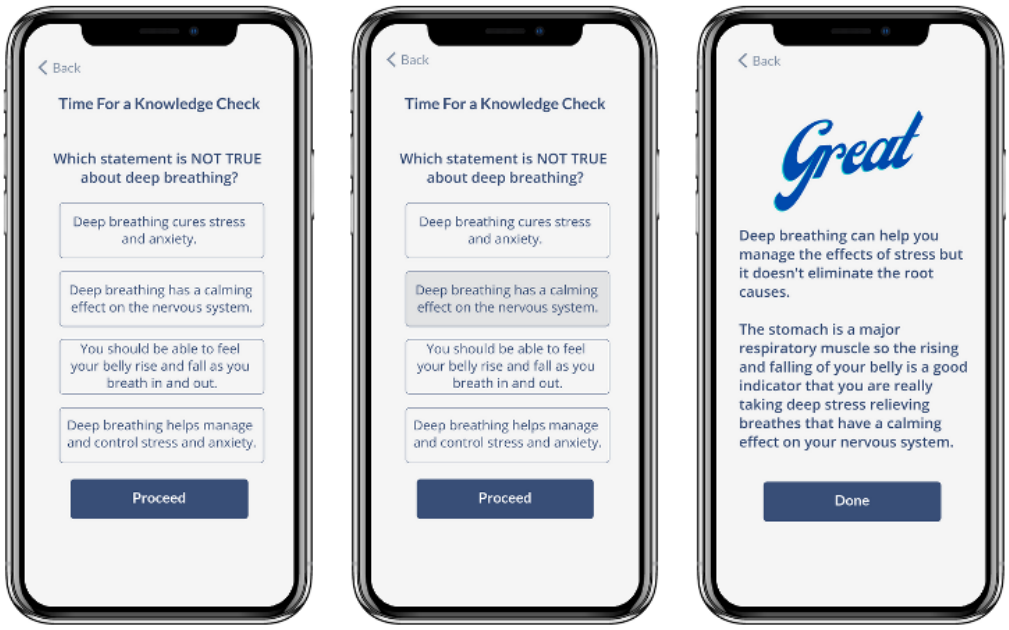


Figure 3. Sample of a knowledge check.



Methodology for Heuristic Evaluation

Nielsen's 10 principles of heuristic evaluation [36] were used for the initial testing of the prototype (Textbox 1). The prototype was given to 22 undergraduate students at a Midwestern university taking a human-computer interaction course in the Spring of 2022 who were trained in conducting heuristic evaluation. No supplemental demographic data were gathered. They were given the WeCanManage prototype during a class period of 1 hour 15 minutes. During the session, students were split into 6 groups, and each group was given 5 tasks to complete using the prototype. We created 3 sets of 5 tasks, and therefore every 2 groups completed the same tasks. The tasks included going through the introduction course module, switching to text and video fields, and filtering the users by a specific disability through the C2P page. Students logged in to classroom computers and accessed Maze, an online testing platform used to monitor assessment details [41], recorded the path taken by students to complete tasks, and presented questions about their experience to help track their progress. At the end of the session, the groups documented violations of the 10 heuristic principles and rated their usability severity on a 0-4 scale, where 0 is not a usability problem and 4 is a usability catastrophe. Furthermore, the student evaluators filled out a questionnaire through Maze

XSL•FO
RenderX

providing feedback and thoughts on the prototype's design. The questionnaire covered their likes and dislikes of the design, their impressions of course modules, and the ease of changing the format of the content.

Textbox 1. Ten principles of heuristic evaluation from Nielsen [36].

1. Visibility of system status
2. Match between system and the real world
3. User control and freedom
4. Consistency and standards
5. Error prevention
6. Recognition rather than recall
7. Flexibility and efficiency of use
8. Aesthetic and minimalist design
9. Help users recognize, diagnose, and recover from errors
10. Help and documentation

Methodology for Usability Testing

We modified the prototype based on the feedback from heuristic evaluation and conducted usability testing over Zoom. We used purposive sampling with targeted outreach through cancer survivorship networks, including both clinical and community. To be eligible for participation, individuals had to meet the following inclusion criteria: be 18 years or older; have a history of breast cancer, head and neck cancer, or sarcoma; have completed active treatment; self-identify as a person with a disability; and possess the ability to understand and communicate in English. Participants received a gift card for their time. Sessions lasted approximately 90 minutes. Sessions were recorded and participants shared their screens for data collection. Participants were told to connect to Zoom on a computer or laptop device. Usability testing occurred between September 2022 and February 2023. As we encountered minor issues with the Maze platform during the heuristic evaluation, including audio malfunctions, we transitioned to Ballpark, an extension of Marvel that facilitated usability testing of the prototype. Participants were given 8 tasks to complete (see Textbox 2). They were told that they were on day 6 of the 4-week period. Consequently, they could access content from sessions 1-6, while subsequent sessions remained locked to replicate the user's sequential navigation experience, with new content being unlocked on a daily basis. The first 6 tasks were based on the course sessions and navigating through each course by reading the content cards and doing related engagement activities. Task 2 required participants to switch the viewing mode using the accessibility features (eye symbol) to the text-only mode, while task 6 involved watching a 1 minute 20 second–long mindfulness video, instead of the default card format. The final 2 tasks (tasks 7 and 8) focused on navigating the Community Forum and C2P sections. After each task, participants rated their satisfaction level and the time taken to complete each task using a 7-point Likert scale. On finishing all 8 tasks, participants had the opportunity to freely explore the app using a "think aloud" approach to express their thoughts and experiences.

To evaluate usability, participants completed the modified SUS, a reliable and valid 10-item questionnaire that assesses usability [42,43]. While the SUS has been around since 1986, it has been shown to be effective in evaluating the usability of recent health apps [44]. To calculate SUS scores, 1 is subtracted from the raw score of the odd-numbered items (those items phrased in a positive way), and the raw score of the even-numbered items (those items phrased in a negative way) is subtracted from 5. The total scores are then multiplied by 2.5 to derive the "standardized SUS score," which ranges from 0 to 100. A SUS score of 68 is considered average usability [45], while a score above 80.3 is deemed an A grade, placing it in the top 10% of scores [46] and corresponding to a narrative rating of good-excellent [47]. In addition, we included open-ended questions to gather feedback on participants' preferences and areas for improvement regarding the app. Examples of these questions include "How easy or difficult was it to see all the content on the screen?" and "What did you think of the design of the course modules?"

To assess the effectiveness of the app design, following a similar approach to Adler et al [48], we evaluated task completion by having 2 independent coders review each recording and code whether the participants

- Completed the task quickly on their own (C)
- Completed the task on their own though it took a little longer (L)
- Needed help to complete the task (H)

The coders achieved an agreement percentage of 87.5%. Any discrepancies were resolved through discussion. To assess efficiency, we analyzed the number of misclicks (clicks outside of clickable areas in the prototype) and the time taken to complete each task.

XSL•FO
RenderX

Textbox 2. Eight tasks given to usability testing participants.

Course

1. Go to the Course and click on the WeCanRelate session. Read through all of the cards.
2. Go to the Course and click on the Introduction session. Switch to Text view to read all the cards at once using the eye symbol on the bottom left of the first screen of the module.
3. Go to the Course and click on the Celebrating & Taking Stock session. Read through all the cards and then go to the reflection. Start "typing" your reflection and post it. Do you see your post accurately reflected?
4. Go to the Course and click on the Straight Talk About Symptoms session. Read through the cards and follow the link to the library and the Understanding the Cancer Rehabilitation Team Fact Sheet.
5. Go to the Course and click on the Deep Breathing session. Read through the content and complete the knowledge check. Did you get the correct answer?
6. Go to the Course and click on the Body Awareness session and go through to the end of the module by watching the video.

Community

1. Go to the Community Forum. Create a new post in the Open Discussion forum. Enter a title, select the community tag, enter text, and post your response.

Connect to Peers

1. Find the Connect to Peers (C2P) option and filter to narrow the search to people who are deaf or hard of hearing.

Ethics Approval

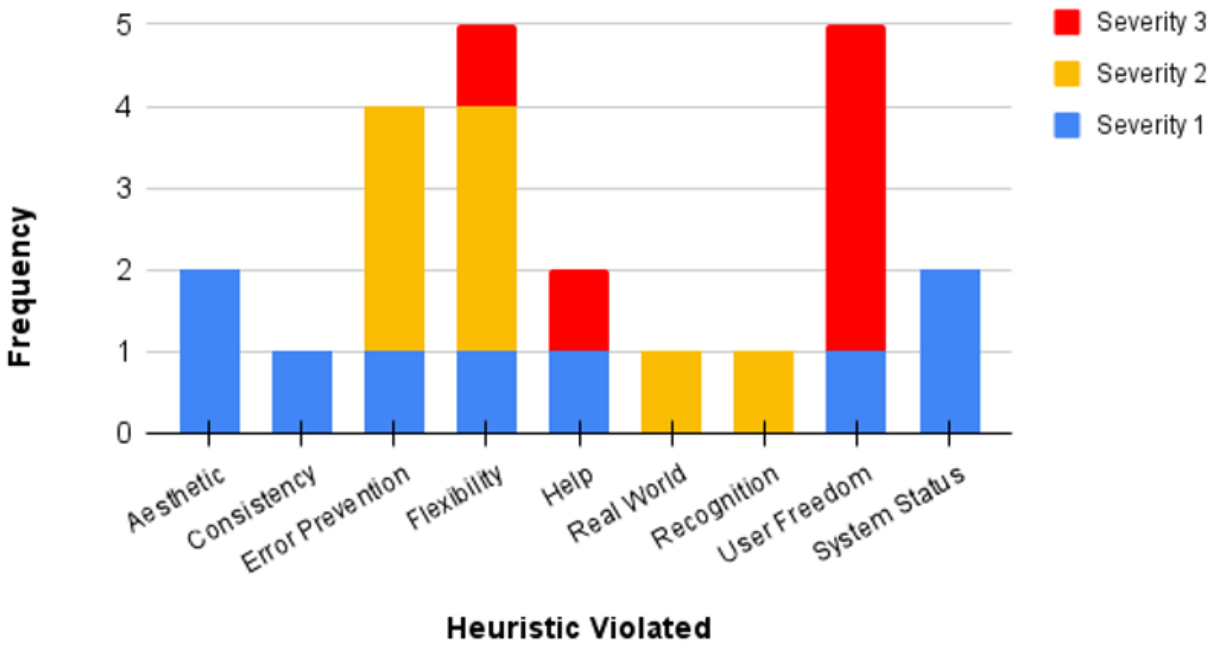
We obtained institutional review board approval from the participating universities in the project (University of Illinois Chicago #2020-1067, Northeastern Illinois University #79, and Northwestern University #NUUIC21CC03).

Results

Results of Heuristic Evaluation

We conducted an analysis of the identified heuristic violations and their severity. The highest severity rating recorded was a 3, as illustrated in Figure 4. The most frequent heuristic violations were related to flexibility, user control, and freedom, followed by error prevention. The issues identified were primarily navigation problems within the prototype, missing back buttons, and font size being too small. Suggestions for improvement were also raised, such as adding an FAQ page, a way to contact the creators or administrators, and including a walk-through or how-to page. Student evaluators expressed appreciation for the images and content, the knowledge check feature, the color scheme, and the layout. They found the app easy to read and navigate. The dislikes expressed included the absence of a help guide and nonfunctional back buttons. Additionally, some groups reported having difficulty finding the format button to switch the mode of learning to text-only or audio.

Figure 4. Graph displaying the frequency and severity of heuristic violations.

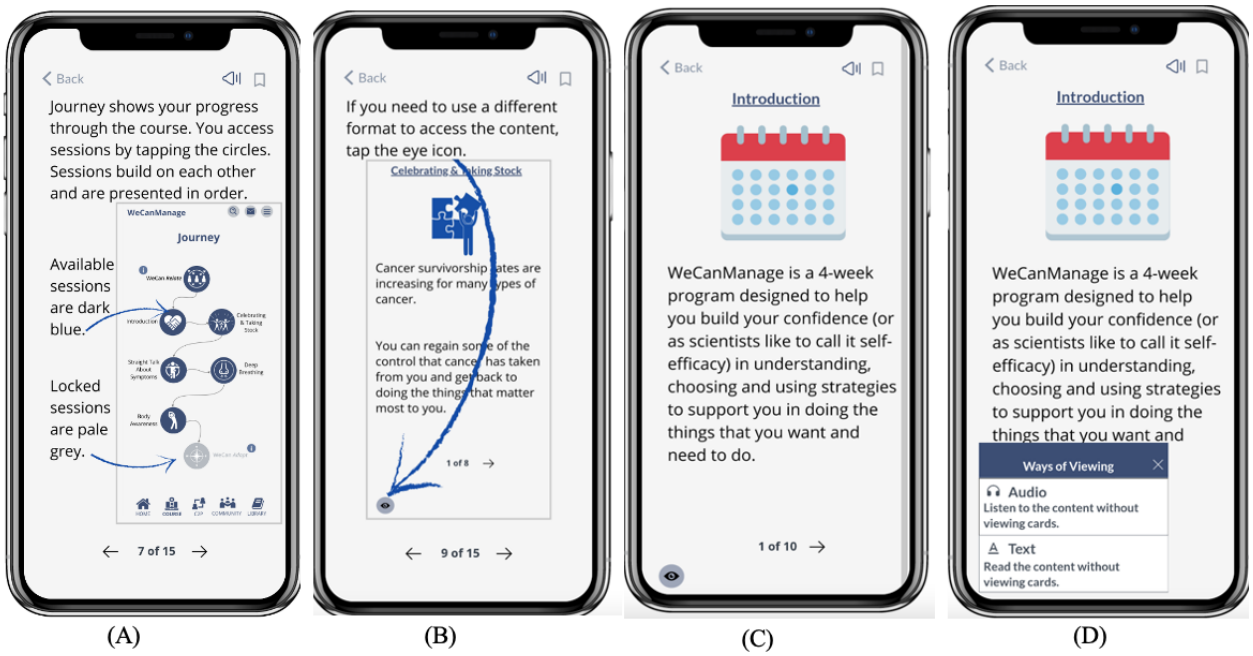


XSL•FO
RenderX

Modifications Based on Heuristic Evaluation

Drawing from the findings of the heuristic evaluation, we enhanced the prototype by introducing a help guide (Figure 5A and B) and seamlessly integrating it into the first course session. We also revised the method for switching accessibility format features (Figure 5C and D). Furthermore, we increased the font size on multiple screens and improved navigation by implementing additional back buttons for a smoother user experience.

Figure 5. Updated prototype screens after heuristic evaluation. (A,B) Help guide incorporated into the first course session. (C,D) Updates to the accessibility format and switching from card view to audio or text views.



Results of Usability Testing

We had 10 cancer survivors with disabilities (9 female, 1 male; 9 White or Caucasian, 1 Black or African American) who completed usability testing. The average age of the participants was 59 years. Usability scores show that participants had an overall positive reception to the design of the prototype. We had an average SUS score of 81; our prototype's usability is therefore considered good to excellent with a grade of an A and in the top 10%.

We assessed participants' satisfaction levels and the time taken to complete each task. The average scores for these 2 measurements are presented in Table 1. Generally, participants exhibited high satisfaction rates; however, lower numbers were observed for task 2 (finding the eye icon to change the accessibility format), task 7 (creating a post in the Community Forum), and task 8 (using the filter in C2P).

In addition, we evaluated the effectiveness of the app design by categorizing participants' task completion into 3 groups: completed quickly (C), completed with a little more time (L), or required assistance to complete the task (H). Overall, most participants completed their tasks without any issues, with only 17 of 80 cases (21%) needing help to complete them (see Figure 6). During task 1, a slight learning curve was observed as some participants had difficulty locating the correct module, leading to the need for assistance in completing the task. However, this issue was not prevalent in subsequent tasks. Task 2 revealed that some participants encountered challenges while switching the card format to text view using the eye symbol, as they had trouble locating the button. In task 4, some participants faced difficulties clicking on the correct resource within the Library as directed in the learning module. For tasks 7 and 8, several participants struggled to navigate both the Community and C2P sections because certain text and icons were too small or unclear in their function, leading to confusion on what to do.

Likewise, while analyzing efficiency based on the number of misclicks per task, tasks 7 and 8 exhibited notably higher misclick rates (Table 2). The table also presents the actual time taken per task, with task 1 showing higher time than the other tasks. As mentioned earlier, task 1 had a learning curve, but it also involved reading the most cards (15 cards) as we integrated the help guide into the first course session. Therefore, this finding is expected given the additional content to review in task 1.

The prototype's help guide received a positive response, with 8 of 10 participants (80%) rating it as very helpful or extremely helpful. Similarly, 8 of 10 participants (80%) reported finding the eye symbol (to change the course format) easily. In response to open-ended questions, participants expressed their likes and dislikes of the prototype and its design. Many participants shared positive opinions on the design and content of the modules, finding them helpful and insightful. The video located within one of the modules received positive feedback, with some expressing a desire for additional videos. The purpose of the Community section was well liked as participants enjoyed

XSL•FO
RenderX

having a place to freely express themselves with other cancer survivors and appreciated the opportunity for users to support each other. The Library resources were found to be informative and useful, covering a wide range of topics.

Our findings were overwhelmingly positive, supported by quotes from participants (some written and some oral):

> *I want to see the whole thing work! I know that this is a prototype, but I want to see more!*
>
> *Great app, it would have been very helpful to me when I was just out [of] treatment.*
>
> *Even though I'm not very comfortable with technology, and that might be because of my age, … I don't think that this would be difficult for me. I think there'd be a real fast learning curve. I felt good and positive when I realized I had learned something, and I could just click on it now without having to think about it.*
>
> *I do like the app. I like that I know I'm not alone feeling this way.*

These participant quotes reflect their enthusiasm and positive experiences with the app, highlighting its potential benefits and ease of use.

On the basis of our session observations and participants' feedback on areas for improvement, we identified several issues:

- Accessibility concerns, including small font sizes and icons, particularly with the navigation arrows on cards, the top navigation bar, and the eye icon.
- Some participants experienced confusion while navigating the Community page when creating new posts.
- Difficulty in locating and using the filter option within the C2P page.
- Participants expressed a desire for an easy way to return to the help guide.
- Feedback indicated a preference for changing the robotic voices used in the audio format for the modules. The prototype used Google US English from voicegenerator.io, but the intention is to have a real person's voice in future implementations.

Addressing these areas for improvement can further enhance the app's usability and user experience.

Table 1. Average satisfaction per task and time per task (out of 7).

Task	Average task satisfaction	Average time satisfaction
1	6.5	6.4
2	5.7	5.9
3	6.6	6.5
4	6.5	6.2
5	6.7	6.3
6	6.8	6.6
7	5.2	5.5
8	5.8	5.7

XSL•FO
RenderX

Figure 6. Graph displaying the frequency of H (required assistance to complete the task), C (completed quickly), and L (completed with a little more time) ratings given to participants as they completed a task.

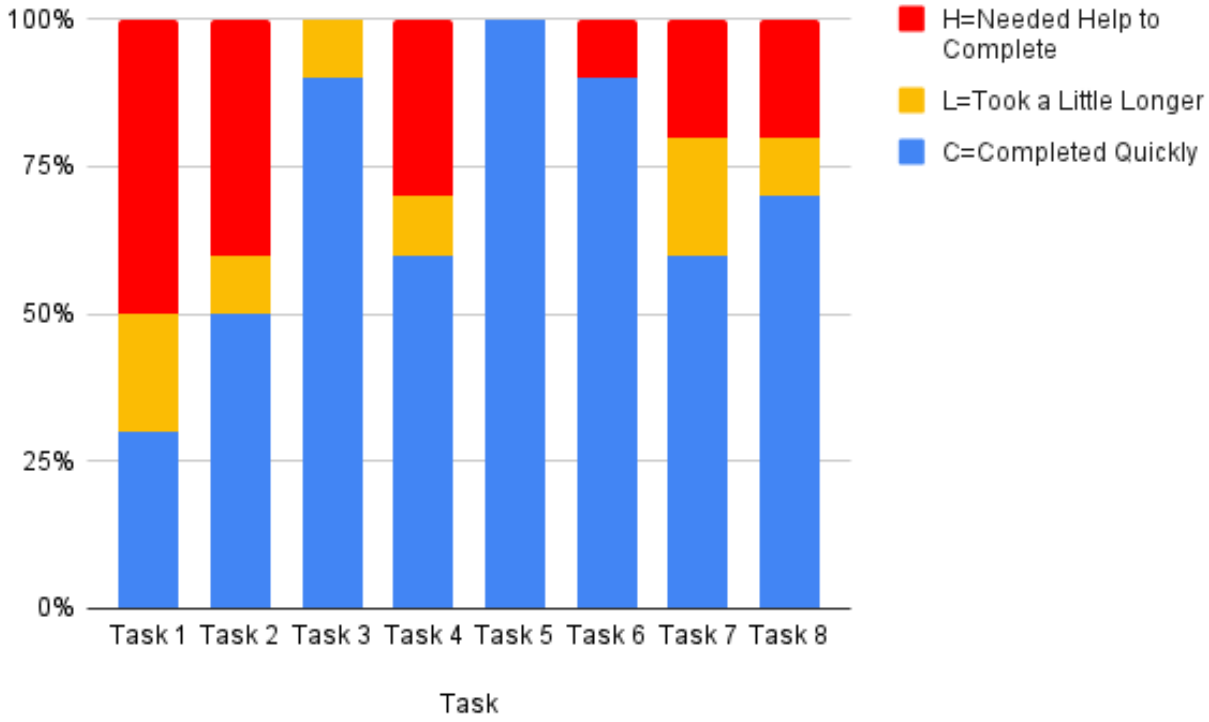


Table 2. Percentage of misclicks and time per task.

Task	Misclicks (%)	Time (minutes)
1	8	3:28
2	4.75	2:28
3	3.30	2:19
4	5.64	2:13
5	0.83	2:15
6	0	1:57
7	19.24	1:34
8	16.38	0:44

Modifications Based on Usability Testing

On the basis of the findings from usability testing, we made several modifications to the prototype. To enhance usability, we increased the sizes of navigation icons, the eye icon, arrows within cards, and the top navigation bar. Throughout the application, we enlarged or bolded fonts for easier reading, including the “create new post” button in the Community section. We redesigned the layout of the Community Forum, increasing text and margins to achieve a cleaner and more concise design. Additionally, we revamped the subscribe button to reduce confusion (see Figures 7 and 8). To improve accessibility, we enlarged the C2P filter. Finally, we added a convenient way to return to the help guide by including it in the hamburger menu icon on the main page. These changes aim to enhance user experience and address the identified issues during usability testing.

XSL•FO
RenderX

Figure 7. Modifications made to the Community before and after usability testing.

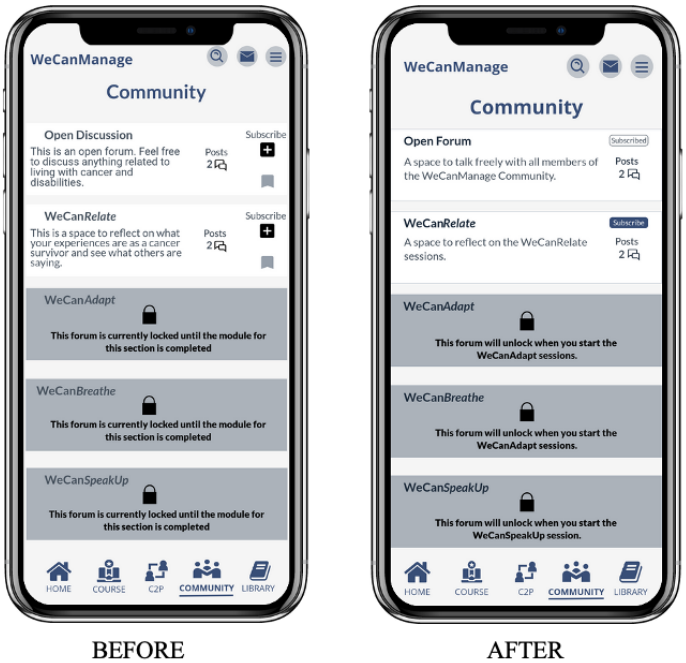
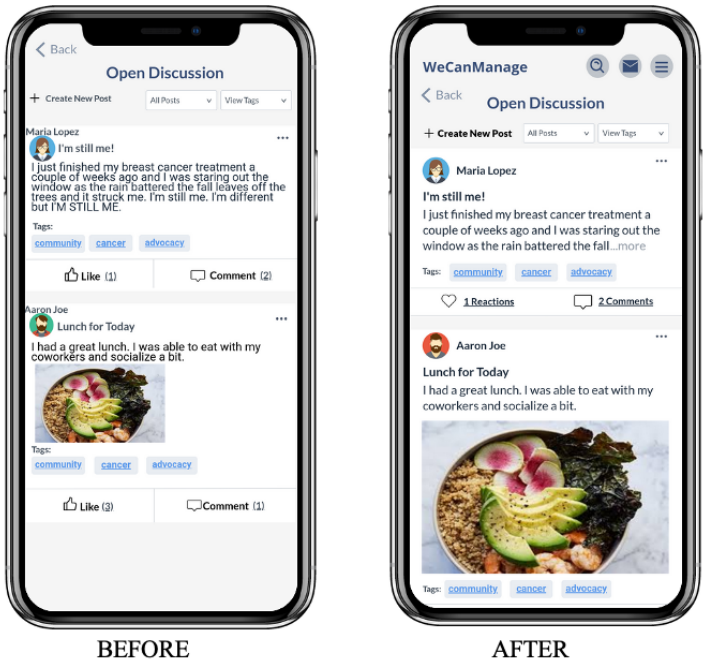


Figure 8. Modifications made to the Open Discussion design before and after usability testing .



Discussion

Principal Findings

Cancer and its treatments can lead to long-term disabilities, significantly impacting a survivor's overall quality of life [10]. Unfortunately, postcancer treatment resources are often limited, further exacerbating the challenges faced by survivors [49,50]. To address this, we developed a high-fidelity prototype for an mHealth app called WeCanManage, aimed at empowering cancer survivors with disabilities to effectively self-manage the long-term effects of cancer treatment. Through conducting the heuristic evaluation, valuable improvements were made, including the incorporation of a helpful guide and the enhancement of accessibility formatting options, ultimately enhancing the overall user experience of the app.

In usability testing, we engaged cancer survivors with disabilities, using multiple methods such as task completion, think aloud strategies, SUS, perceived task satisfaction, and

XSL•FO
RenderX

open-ended questions. These methods have been extensively used to evaluate various applications, with the SUS being one of the commonly used questionnaires [21]. The results of usability testing were overwhelmingly positive, with cancer survivors expressing appreciation for the app's content, features, and design. The prototype achieved an impressive SUS score of 81, ranking it in the top 10% of scores and earning an A grade. Moreover, participants reported high satisfaction levels and efficiency, with average scores of 6.2 and 6.1 (out of 7), respectively. Conducting usability testing enabled us to thoroughly assess the app's overall effectiveness, efficiency, satisfaction, and usability. We were able to identify areas for improvement, particularly in terms of accessibility. The insights gained from this testing process have allowed us to refine and enhance the app, ensuring a positive user experience for cancer survivors with disabilities.

In a study by Fuller-Tyszkiewicz et al [24], end users rated an mHealth prototype higher in usability and reported a more positive experience than clinical experts. Interestingly, users did not share the same concerns about the amount and layout of content presented as the experts had anticipated [24]. This discrepancy underscores the significance of testing potential users to tailor the app to their specific needs and preferences. While expert opinions (whether clinical or in design) are valuable, evaluating an app on actual users is ideal.

Implications for Designers and Researchers

One of our primary findings is the importance of accessibility when designing applications for cancer survivors. Our app was specifically designed for cancer survivors with disabilities, and as such, we incorporated customized options to switch the learning style. Users could choose between clicking through content cards and accessing audio or text-only views. This flexibility proved to be helpful, particularly for participants with cognitive issues like "chemo brain," who found it easier to navigate the audio versions of the course sessions. However, during testing, we identified other accessibility concerns related to font sizes and icons. Some users found them too small to see, click on, and navigate effectively. Addressing these issues is essential to ensure an inclusive and user-friendly experience for all app users.

The importance of having a help feature was revealed during heuristic evaluation, and through usability testing, we learned that users expressed a desire for a convenient way to return to the help guide. In response to this feedback, we have now incorporated the option to access the help guide directly from our main menu.

One comment expressed by many of our participants was how lonely the experience of a cancer survivor is. Consistent with findings from other studies that highlight the significance of social features in mHealth apps [51], participants expressed their appreciation for the Community Forum and C2P sections. These features provide a valuable opportunity for them to connect with others facing similar situations, fostering a sense of community and support. Additionally, participants reported that reading the content in the course sessions made them realize that their experiences were shared by others, helping them feel less isolated and reassured that they were not alone in their journey. When asked what they liked about the app, one participant wrote the following: "The information, reliable and trustworthy, … and the realization that I am not alone."

Limitations

Our aim was to achieve a minimum of 12 participants for usability testing, as SUS results are ideally derived from 12 or more participants [52,53]. However, we encountered challenges in recruitment because of technical difficulties, such as some participants lacking access to a laptop or facing issues with Zoom and screen sharing, leading to incomplete usability testing. Additionally, recruitment was hindered by our specific inclusion criteria, which focused on individuals who identified as having a disability. These challenges impacted our ability to reach the desired number of participants for the usability testing phase. Nevertheless, it is worth noting that according to Nielsen [54], 5 participants are typically adequate for identifying usability problems. Thus, we can reasonably infer that our processes have successfully identified the majority of issues, providing a level of confidence in the validity of our findings despite the lower number of participants in the usability testing phase. Additionally, it is worth mentioning that several studies evaluating mHealth prototypes have used the SUS with fewer than 12 participants [29,31,37]. We encountered instances where some participants experienced lingering effects of cancer and its treatment, but they did not self-identify as having a disability, resulting in their exclusion from usability testing. This finding has important implications for the implementation and adoption of WeCanManage, ensuring that cancer survivors experiencing disabling aftereffects can fully benefit from the tool and appreciate its relevance and value in their daily lives and experiences.

Furthermore, as this was a prototype, not all features were fully implemented (eg, the ability to create a post on the forum or direct message a user was mimicked), which may have caused some participants to encounter difficulties in the Community section of the prototype. In addition, during usability testing, participants expressed concerns regarding text and icon sizes. It is important to note that the testing was conducted over Zoom using computers (not mobile devices), and the prototype's size (matching that of a phone) might have posed challenges during interaction, which may not be representative of the real application's experience. Finally, it is worth noting that the age of participants and their level of comfort with technology might have influenced their overall experience [55]. Nevertheless, because these individuals constitute our target user base, it remains essential for us to maintain the app's usability and accessibility to meet their needs.

Conclusions

When creating an mHealth app, it is crucial to evaluate it with the target users in mind, in our case, cancer survivors with disabilities. Usability testing allowed us to identify the design's strengths and areas requiring improvement. The WeCanManage prototype achieved a SUS score of 81, placing it in the top 10% of scores. Our future work will involve feasibility testing of an implemented web-based mobile app of WeCanManage. This will enable us to further refine the application and ensure that

XSL•FO
RenderX

it meets the needs and preferences of our target users, enhancing its overall usability and impact.

Acknowledgments

This research was supported by the National Cancer Institute of the National Institutes of Health (U54CA202995, U54CA202997, and U54CA203000). The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health. Additional support was provided by Northeastern Illinois University's Student Center for Science Engagement, Committee on Organized Research, and College of Business and Technology. The authors would also like to acknowledge contributions from the following team members: Melissa Delgado, Daniela Raudales Reyes, Elizabeth Jarvis, Sabrina Cadena, and Bruriah Horowitz.

Conflicts of Interest

None declared.

References

1. Cancer facts and figures 2023. American Cancer Society. 2023. URL: https://www.cancer.org/research/cancer-facts-statistics/all-cancer-facts-figures/2023-cancer-facts-figures.html [accessed 2024-02-21]
2. National Research Council. In: Stovall E, Hewitt M, Greenfield S, editors. From Cancer Patient to Cancer Survivor: Lost in Transition. Washington, DC: National Academies Press; 2005.
3. Alfano CM, Ganz PA, Rowland JH, Hahn EE. Cancer survivorship and cancer rehabilitation: revitalizing the link. J Clin Oncol 2012;30(9):904-906 [FREE Full text] [doi: 10.1200/JCO.2011.37.1674] [Medline: 22355063]
4. Henderson F, Cross AJ, Baraniak AR. 'A new normal with chemobrain': experiences of the impact of chemotherapy-related cognitive deficits in long-term breast cancer survivors. Health Psychol Open 2019;6(1):2055102919832234 [FREE Full text] [doi: 10.1177/2055102919832234] [Medline: 30873289]
5. Nguyen NP, Sallah S, Karlsson U, Antoine JE. Combined chemotherapy and radiation therapy for head and neck malignancies: quality of life issues. Cancer 2002;94(4):1131-1141 [FREE Full text] [doi: 10.1002/cncr.10257] [Medline: 11920484]
6. International Classification of Functioning, Disability, and Health: ICF. Geneva: World Health Organization; 2001.
7. What is the definition of disability under the ADA? ADA National Network. 2015. URL: https://adata.org/faq/what-definition-disability-under-ada [accessed 2024-02-21]
8. Winters CA, Marshall HK, Victorson DE, Adler RF, Magasi S. Does the Canadian Occupational Performance Measure (COPM) capture the complex experiences of cancer survivors? A mixed methods approach. Occup Ther Health Care 2023:1-17 [FREE Full text] [doi: 10.1080/07380577.2023.2215875] [Medline: 37302410]
9. Magasi S, Marshall HK, Winters C, Victorson D. Cancer survivors' disability experiences and identities: a qualitative exploration to advance cancer equity. Int J Environ Res Public Health 2022;19(5):3112 [FREE Full text] [doi: 10.3390/ijerph19053112] [Medline: 35270802]
10. Newman RM, Alfano CM, Radomski MV, Pergolotti M, Wolf TJ, Sleight AG, et al. Catalyzing research to optimize cancer survivors' participation in work and life roles. OTJR (Thorofare N J) 2019;39(4):189-196 [FREE Full text] [doi: 10.1177/1539449219844749] [Medline: 31046601]
11. Raj VS, Pugh TM, Yaguda SI, Mitchell CH, Mullan SS, Garces NS. The who, what, why, when, where, and how of team-based interdisciplinary cancer rehabilitation. Semin Oncol Nurs 2020;36(1):150974 [FREE Full text] [doi: 10.1016/j.soncn.2019.150974] [Medline: 31955923]
12. Stubblefield MD. The underutilization of rehabilitation to treat physical impairments in breast cancer survivors. PM R 2017;9(9S2):S317-S323 [FREE Full text] [doi: 10.1016/j.pmrj.2017.05.010] [Medline: 28942906]
13. What is mobile technology? IBM. 2023. URL: https://www.ibm.com/topics/mobile-technology [accessed 2023-05-23]
14. Mobile fact sheet. Pew Research Center. 2021. URL: https://www.pewresearch.org/internet/fact-sheet/mobile/ [accessed 2023-05-21]
15. Curigliano G, Banerjee S, Cervantes A, Garassino MC, Garrido P, Girard N, et al. Managing cancer patients during the COVID-19 pandemic: an ESMO multidisciplinary expert consensus. Ann Oncol 2020;31(10):1320-1335 [FREE Full text] [doi: 10.1016/j.annonc.2020.07.010] [Medline: 32745693]
16. Pohl M. 325,000 mobile health apps available in 2017—Android now the leading mHealth platform. Research2Guidance. 2017. URL: https://research2guidance.com/325000-mobile-health-apps-available-in-2017/ [accessed 2024-02-21]
17. Charbonneau DH, Hightower S, Katz A, Zhang K, Abrams J, Senft N, et al. Smartphone apps for cancer: a content analysis of the digital health marketplace. Digit Health 2020;6:2055207620905413 [FREE Full text] [doi: 10.1177/2055207620905413] [Medline: 32110428]
18. Boulley GE, Leroy T, Bernetière C, Paquienseguy F, Desfriches-Doria O, Préau M. Digital health interventions to help living with cancer: a systematic review of participants' engagement and psychosocial effects. Psychooncology 2018;27(12):2677-2686 [FREE Full text] [doi: 10.1002/pon.4867] [Medline: 30152074]

XSL•FO
RenderX

19. Silva EH, Lawler S, Langbecker D. The effectiveness of mHealth for self-management in improving pain, psychological distress, fatigue, and sleep in cancer survivors: a systematic review. J Cancer Surviv 2019;13(1):97-107 [FREE Full text] [doi: 10.1007/s11764-018-0730-8] [Medline: 30635865]
20. Zheng C, Chen X, Weng L, Guo L, Xu H, Lin M, et al. Benefits of mobile apps for cancer pain management: systematic review. JMIR Mhealth Uhealth 2020;8(1):e17055 [FREE Full text] [doi: 10.2196/17055] [Medline: 32012088]
21. Maramba I, Chatterjee A, Newman C. Methods of usability testing in the development of eHealth applications: a scoping review. Int J Med Inform 2019;126:95-104 [FREE Full text] [doi: 10.1016/j.ijmedinf.2019.03.018] [Medline: 31029270]
22. Hartzler AL, Venkatakrishnan A, Mohan S, Silva M, Lozano P, Ralston JD, et al. Acceptability of a team-based mobile health (mHealth) application for lifestyle self-management in individuals with chronic illnesses. Annu Int Conf IEEE Eng Med Biol Soc 2016;2016:3277-3281 [FREE Full text] [doi: 10.1109/EMBC.2016.7591428] [Medline: 28269007]
23. Martinez W, Threatt AL, Rosenbloom ST, Wallston KA, Hickson GB, Elasy TA. A patient-facing diabetes dashboard embedded in a patient web portal: design sprint and usability testing. JMIR Hum Factors 2018;5(3):e26 [FREE Full text] [doi: 10.2196/humanfactors.9569] [Medline: 30249579]
24. Fuller-Tyszkiewicz M, Richardson B, Klein B, Skouteris H, Christensen H, Austin D, et al. A mobile app-based intervention for depression: end-user and expert usability testing study. JMIR Ment Health 2018;5(3):e54 [FREE Full text] [doi: 10.2196/mental.9445] [Medline: 30139722]
25. Santesteban-Echarri O, Tang J, Fernandes J, Addington J. Development and usability testing of SOMO, a mobile-based application to monitor social functioning for youth at clinical high-risk for psychosis. Digit Psych 2020;1(1):4-19 [FREE Full text] [doi: 10.24989/dp.v1i1.1806]
26. Reynoldson C, Stones C, Allsop M, Gardner P, Bennett MI, Closs SJ, et al. Assessing the quality and usability of smartphone apps for pain self-management. Pain Med 2014;15(6):898-909 [FREE Full text] [doi: 10.1111/pme.12327] [Medline: 24422990]
27. Woods L, Cummings E, Duff J, Walker K. Design thinking for mHealth application co-design to support heart failure self-management. Stud Health Technol Inform 2017;241:97-102. [Medline: 28809190]
28. Fu MR, Axelrod D, Guth AA, Rampertaap K, El-Shammaa N, Hiotis K, et al. mHealth self-care interventions: managing symptoms following breast cancer treatment. Mhealth 2016;2(28):1-18 [FREE Full text] [doi: 10.21037/mhealth.2016.07.03] [Medline: 27493951]
29. Donald M, Beanlands H, Straus SE, Smekal M, Gil S, Elliott MJ, et al. A web-based self-management support prototype for adults with chronic kidney disease (My kidneys My Health): co-design and usability testing. JMIR Form Res 2021;5(2):e22220 [FREE Full text] [doi: 10.2196/22220] [Medline: 33560245]
30. Ospina-Pinillos L, Davenport T, Diaz AM, Navarro-Mancilla A, Scott EM, Hickie IB. Using participatory design methodologies to co-design and culturally adapt the Spanish version of the Mental Health eClinic: qualitative study. J Med Internet Res 2019;21(8):e14127 [FREE Full text] [doi: 10.2196/14127] [Medline: 31376271]
31. Harte R, Quinlan LR, Glynn L, Rodríguez-Molinero A, Baker PM, Scharf T, et al. Human-centered design study: enhancing the usability of a mobile phone app in an integrated falls risk detection system for use by older adult users. JMIR Mhealth Uhealth 2017;5(5):e71 [FREE Full text] [doi: 10.2196/mhealth.7046] [Medline: 28559227]
32. Schnall R, Rojas M, Bakken S, Brown W, Carballo-Dieguez A, Carry M, et al. A user-centered model for designing consumer mobile health (mHealth) applications (apps). J Biomed Inform 2016;60:243-251 [FREE Full text] [doi: 10.1016/j.jbi.2016.02.002] [Medline: 26903153]
33. Iacobelli F, Adler RF, Buitrago D, Buscemi J, Corden ME, Perez-Tamayo A, et al. Designing an mHealth application to bridge health disparities in Latina breast cancer survivors: a community-supported design approach. Design Health (Abingdon) 2018;2(1):58-76 [FREE Full text] [doi: 10.1080/24735132.2018.1452871] [Medline: 30506017]
34. Beatty L, Koczwara B, Butow P, Turner J, Girgis A, Schofield P, et al. Development and usability testing of a web-based psychosocial intervention for women living with metastatic breast cancer: finding my way-advanced. J Cancer Surviv 2021;15(3):403-409 [FREE Full text] [doi: 10.1007/s11764-021-01019-5] [Medline: 33723741]
35. Børøsund E, Mirkovic J, Clark MM, Ehlers SL, Andrykowski MA, Bergland A, et al. A stress management app intervention for cancer survivors: design, development, and usability testing. JMIR Form Res 2018;2(2):e19 [FREE Full text] [doi: 10.2196/formative.9954] [Medline: 30684438]
36. Nielsen J. Heuristic evaluation. In: Nielsen J, Mack RL, editors. Usability Inspection Methods. New York: John Wiley & Sons; 1994.
37. Derks YP, Klaassen R, Westerhof GJ, Bohlmeijer ET, Noordzij ML. Development of an ambulatory biofeedback app to enhance emotional awareness in patients with borderline personality disorder: multicycle usability testing study. JMIR Mhealth Uhealth 2019;7(10):e13479 [FREE Full text] [doi: 10.2196/13479] [Medline: 31617851]
38. Magasi S, Banas J, Horowitz B, Reis JP, The K, Wilson T, et al. WeCanConnect: development of a community-informed mhealth tool for people with disabilities and cancer. Prog Community Health Partnersh 2019;13(5):49-59 [FREE Full text] [doi: 10.1353/cpr.2019.0038] [Medline: 31378735]
39. Rapid prototyping, testing and handoff for modern design teams. Marvel. URL: https://marvelapp.com/ [accessed 2023-09-20]
40. Adler RF, Morales P, Sotelo J, Magasi S. Developing an mHealth app for empowering cancer survivors with disabilities: co-design study. JMIR Form Res 2022;6(7):e37706 [FREE Full text] [doi: 10.2196/37706] [Medline: 35881439]

XSL•FO
RenderX

41. The holistic solution for continuous product discovery. Maze. URL: https://maze.co/ [accessed 2023-09-20]
42. Brooke J. SUS: a quick and dirty usability scale. In: Thomas B, Weerdmeester B, McClelland IL, Jordan J, editors. Usability Evaluation In Industry. London, Bristol, PA: CRC Press; 1996:189-194.
43. Bangor A, Kortum PT, Miller JT. An empirical evaluation of the system usability scale. Intl J Human–Computer Interact 2008;24(6):574-594 [FREE Full text] [doi: 10.1080/10447310802205776]
44. Hyzy M, Bond R, Mulvenna M, Bai L, Dix A, Leigh S, et al. System usability scale benchmarking for digital health apps: meta-analysis. JMIR Mhealth Uhealth 2022;10(8):e37290 [FREE Full text] [doi: 10.2196/37290] [Medline: 35980732]
45. Sauro J, Lewis JR. Quantifying the User Experience: Practical Statistics for User Research. Cambridge, MA: Morgan Kaufmann; 2016.
46. Sauro J. Measuring usability with the System Usability Scale (SUS). MeasuringU. 2011. URL: https://measuringu.com/sus/ [accessed 2023-08-01]
47. Bangor A, Kortum P, Miller J. Determining what individual SUS scores mean: adding an adjective rating scale. J usability Stud 2009;4(3):114-123.
48. Adler RF, Paley A, Li Zhao AL, Pack H, Servantez S, Pah AR, et al. A user-centered approach to developing an AI system analyzing U.S. federal court data. Artif Intell Law 2022;31(3):547-570 [FREE Full text] [doi: 10.1007/s10506-022-09320-z]
49. Barsevick AM, Leader A, Bradley PK, Avery T, Dean LT, DiCarlo M, et al. Post-treatment problems of African American breast cancer survivors. Support Care Cancer 2016;24(12):4979-4986 [FREE Full text] [doi: 10.1007/s00520-016-3359-z] [Medline: 27543161]
50. Ridner SH, Rhoten BA, Radina ME, Adair M, Bush-Foster S, Sinclair V. Breast cancer survivors' perspectives of critical lymphedema self-care support needs. Support Care Cancer 2016;24(6):2743-2750 [FREE Full text] [doi: 10.1007/s00520-016-3089-2] [Medline: 26810422]
51. Jessen S, Mirkovic J, Ruland CM. Creating gameful design in mHealth: a participatory co-design approach. JMIR Mhealth Uhealth 2018;6(12):e11579 [FREE Full text] [doi: 10.2196/11579] [Medline: 30552080]
52. Lewis JR. The system usability scale: past, present, and future. Int J Hum-Comput Interact 2018;34(7):577-590 [FREE Full text]
53. Tullis TS, Stetson JN. A comparison of questionnaires for assessing website usability. 2004 Presented at: Usability Professionals' Association Conference 2004: 13th Annual UPA Conference; June 7-11, 2004; Minneapolis, Minnesota, USA. [doi: 10.1007/978-1-84628-795-4_3]
54. Nielsen J. Why you only need to test with 5 users. Nielsen Norman Group. 2000. URL: https://www.nngroup.com/articles/why-you-only-need-to-test-with-5-users/ [accessed 2024-02-21]
55. Iancu I, Iancu B. Designing mobile technology for elderly. A theoretical overview. Technol Forecast Soc Change 2020;155:119977 [FREE Full text] [doi: 10.1016/j.techfore.2020.119977]

Abbreviations

C2P: Connect to Peers
mHealth: mobile health
SUS: System Usability Scale

Edited by C Jacob; submitted 03.08.23; peer-reviewed by E Baker, C Baxter; comments to author 06.09.23; revised version received 07.11.23; accepted 11.11.23; published 02.04.24.

Please cite as:
Adler RF, Baez K, Morales P, Sotelo J, Victorson D, Magasi S
Evaluating the Usability of an mHealth App for Empowering Cancer Survivors With Disabilities: Heuristic Evaluation and Usability Testing
JMIR Hum Factors 2024;11:e51522
URL: https://humanfactors.jmir.org/2024/1/e51522
doi:10.2196/51522
PMID:38564261

©Rachel F Adler, Kevin Baez, Paulina Morales, Jocelyn Sotelo, David Victorson, Susan Magasi. Originally published in JMIR Human Factors (https://humanfactors.jmir.org), 02.04.2024. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on https://humanfactors.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

User Perceptions of Visual Clot in a High-Fidelity Simulation Study: Mixed Qualitative-Quantitative Study

Greta Gasciauskaite[1], MD; Clara Castellucci[1], MD; Amos Malorgio[1], MD; Alexandra D Budowski[1], MD; Giovanna Schweiger[1], MD; Michaela Kolbe[2], Dr rer nat; Daniel Fries[1], MD; Bastian Grande[1], MD; Christoph B Nöthiger[1], MD; Donat R Spahn[1], MD; Tadzio R Roche[1], MD; David W Tscholl[1], MD; Samira Akbas[1], MD

[1]Institute of Anesthesiology, University Hospital Zurich, Zurich, Switzerland

[2]Simulation Centre, University Hospital Zurich, Zurich, Switzerland

Corresponding Author:
Samira Akbas, MD
Institute of Anesthesiology
University Hospital Zurich
Raemistrasse 100
Zurich, 8091
Switzerland
Phone: 41 432532242
Email: Samira.Akbas@usz.ch

Abstract

Background: Viscoelastic hemostatic assays, such as rotational thromboelastometry (ROTEM) or thromboelastography, enable prompt diagnosis and accelerate targeted treatment. However, the complex interpretation of the results remains challenging. Visual Clot—a situation awareness-based visualization technology—was developed to assist clinicians in interpreting viscoelastic tests.

Objective: Following a previous high-fidelity simulation study, we analyzed users' perceptions of the technology, to identify its strengths and limitations from clinicians' perspectives.

Methods: This is a mixed qualitative-quantitative study consisting of interviews and a survey. After solving coagulation scenarios using Visual Clot in high-fidelity simulations, we interviewed anesthesia personnel about the perceived advantages and disadvantages of the new tool. We used a template approach to identify dominant themes in interview responses. From these themes, we defined 5 statements, which were then rated on Likert scales in a questionnaire.

Results: We interviewed 77 participants and 23 completed the survey. We identified 9 frequently mentioned topics by analyzing the interview responses. The most common themes were "positive design features," "intuitive and easy to learn," and "lack of a quantitative component." In the survey, 21 respondents agreed that Visual Clot is easy to learn and 16 respondents stated that a combination of Visual Clot and ROTEM would help them manage complex hemostatic situations.

Conclusions: A group of anesthesia care providers found Visual Clot well-designed, intuitive, and easy to learn. Participants highlighted its usefulness in emergencies, especially for clinicians inexperienced in coagulation management. However, the lack of quantitative information is an area for improvement.

(JMIR Hum Factors 2024;11:e47991) doi:10.2196/47991

KEYWORDS

Visual Clot; blood coagulation; blood coagulation test; hemostasis; rotational thromboelastometry; situation awareness; user-centered design; visualization; user; perception; interpretation; thromboelastography; viscoelastic hemostatic; technology; coagulation; quantitative information

Introduction

Rapid hemostatic assessment is essential to targeted coagulation management in acute bleeding [1]. Increasingly used viscoelastic hemostatic assays, such as rotational thromboelastometry (ROTEM) or thromboelastography, enable faster insights into coagulation dysfunction than conventional laboratory tests. Standard coagulation assays are not optimal for managing acute hemorrhages that require rapid therapeutic action, as it often takes more than an hour to obtain the results [2,3]. European

XSL•FO
RenderX

and North American transfusion recommendations underline the advantages of viscoelastic hemostatic assays for managing trauma and severe perioperative bleeding, including the reduced need for transfusions, fewer perioperative complications, shorter hospitalization, and lower overall treatment costs [1,4,5]. Its usefulness has been demonstrated in many operative areas such as obstetrics [6,7], pediatric surgery [8], transplantation [9], cardiac surgery [10,11], neurosurgery [12], and burn surgery [13,14]. Viscoelastic hemostatic tests are also paramount in the diagnostic and treatment adjustment of hematological disorders, such as inherited afibrinogenemia, hemophilia, or multiple myeloma [15-18]. However, despite these technologies' widespread use and considerable advantages, their results' interpretation remains challenging and requires well-trained clinical personnel [19-21]. Visual Clot—a situation awareness-based visualization technology—was developed to support health care professionals in interpreting viscoelastic test results by reducing the complexity of their presentation. Based on raw ROTEM data, the results are displayed in real time as a 3D animated model of a blood clot to represent various elements of hemostasis, including platelets, plasmatic factors, and fibrin. It can also effectively illustrate the influence of heparin and hyperfibrinolysis [22] (Multimedia Appendix 1). In a high-fidelity simulation study, anesthesiologists using Visual Clot were 2.2 times more likely to articulate the correct therapeutic approach. In addition, these anesthesiologists had a lower median time to administer the first appropriate targeted coagulation product. Overall, physicians presented with the results of viscoelastic testing using Visual Clot were approximately 56% more likely to provide accurate therapeutic interventions. In the same study, physicians were 3.5 times more likely to feel confident in their decisions when working with Visual Clot compared to traditional ROTEM results [23]. In the first computer-based study analyzing user perceptions of Visual Clot, participants described the technology as well-designed, easy to learn, and intuitive [24]. The guiding principles of the Visual Clot technology that result in enhanced situation awareness include Endsley's user-centered design principles [25], Wittgenstein's philosophy as articulated in Tractatus Logico-Philosophicus [26], and insights from the National Aeronautics and Space Administration (NASA) publication "On Organization of Information: Approach and Early Work" by Degani et al [27]. Endsley's principles emphasize the use of direct visual representations of data to enhance situational awareness, a central principle in Visual Clot's data visualization. Wittgenstein's theory emphasizes the importance of logical representations that meaningfully correspond to the reality they are intended to represent. Visual Clot follows this principle by visually representing elements such as fibrin, platelets, plasmatic factors, hyperfibrinolysis, and bleeding. Following NASA's approach, Visual Clot strives to achieve the highest level of "order and wholeness" by consolidating all essential data into a single display. The primary goal of Visual Clot technology is to provide the care provider with situational information quickly and with minimal cognitive load. In this study, we aimed to capture and analyze perceptions of anesthesia personnel working with Visual Clot in a high-fidelity simulation to identify the strengths and recognize the potential for future improvements.

Methods

Ethical Considerations

The Cantonal Ethics Committee of the Canton of Zurich reviewed the study protocol and issued a declaration of no objection (Business Management System for Ethics Committees Number Req-2021-01112). Furthermore, each participant gave informed consent to use his or her data for research purposes. Participation was voluntary and without financial compensation.

Study Design

We conducted a researcher-initiated single-center mixed qualitative-quantitative study at the University Hospital Zurich, Institute of Anesthesiology, Switzerland. Study participants were anesthesia personnel, including staff anesthesiologists, residents, and nurses. After participating in a high-fidelity simulation study of perioperative bleeding scenarios, where they worked with Visual Clot and ROTEM, we interviewed participants on their perceptions of Visual Clot technology.

As a second step, the same participants received an email invitation to participate in a survey a few weeks later. They rated statements we generated from identified and frequently mentioned themes in interview responses on a Likert scale.

Previous High-Fidelity Simulation Study

In the high-fidelity simulation study [23], anesthesia teams, composed of a staff anesthesiologist, a resident, and an anesthesia nurse, participated in high-fidelity perioperative bleeding scenarios using either Visual Clot or ROTEM. The primary outcome of the study was correct targeted coagulation therapy. Secondary outcomes were time to targeted coagulation therapy, confidence, and workload.

ROTEM is the standard of care for managing acute hemorrhage in the study center, so all participants were familiar with the technology before participating [20]. Some participants had taken part in previous Visual Clot studies and, therefore, were already familiar with the technology [22,24].

Nevertheless, before the simulations began, we gave a 10-minute presentation that reviewed ROTEM and introduced Visual Clot. Multimedia Appendix 1 provides an instructional video of Visual Clot. Participants were invited to ask questions freely before starting work in the simulation environment. Each team solved 1 of 4 different perioperative bleeding scenarios, which were randomly allocated. We ended the scenarios when all necessary therapeutic measures were derived or, at the latest, after 15 minutes. Figure 1 illustrates an example of a Visual Clot printout used in the simulation study.

XSL•FO
RenderX

Figure 1. A Visual Clot result presentation showing a fibrin deficiency. The fibrin in the clot is shown as a dashed line, indicating its absence. The blood drops indicate the presence of a coagulation pathology.

Participant Interviews

After the simulations, we encouraged the participants to freely verbalize their thoughts in a distraction-free environment while the data collectors made field notes. The only suggestion to the participants before the interviews was to verbalize their positive and negative opinions of Visual Clot. The participants could define final adjustments in the collected answers at the end of the interviews.

Survey

In the second step, we formulated 5 statements to summarize the insights gathered during the interviews. The statements were submitted for evaluation on a 5-point Likert scale graded from "strongly agree" to "strongly disagree." An email invitation was sent to all interviewed participants.

Outcomes and Statistical Analyses

Part I: Participant Interviews

Collected interview responses were translated from original German to English using a translation system DeepL (DeepL GmbH). Multimedia Appendix 2 provides the complete translated field notes.

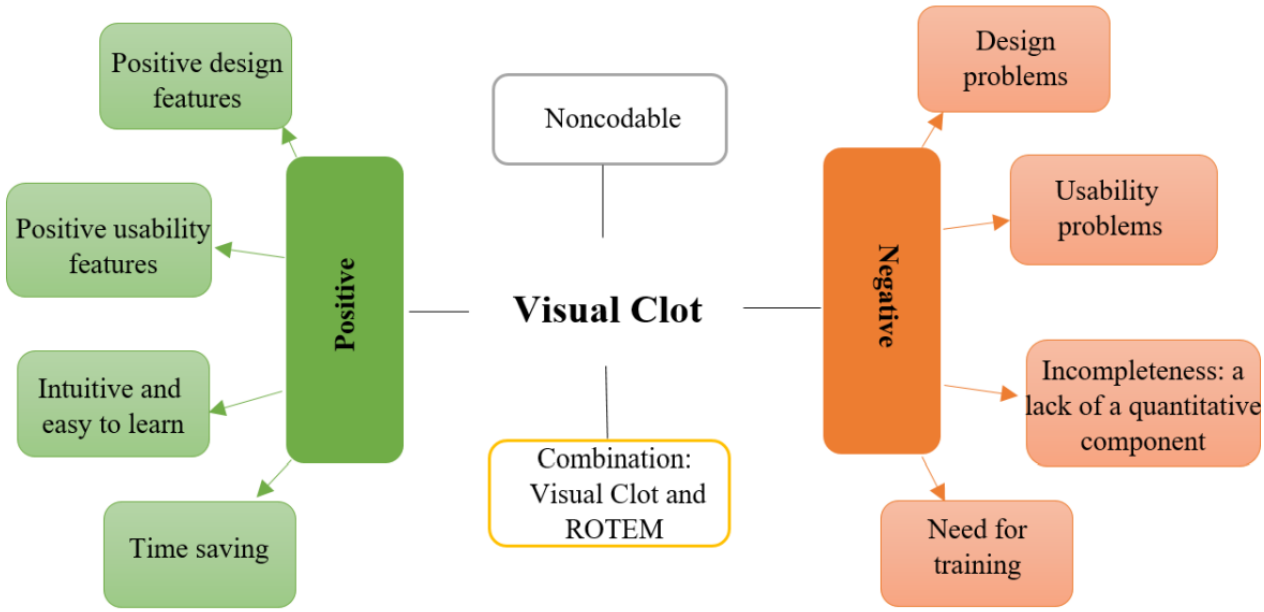
The most commonly used terms in positive and negative responses were identified using the word count function. Word groups with the same root were united, excluding the frequently used filler words such as "to," "and," or "the" (Table 1). Using a template approach [28] we identified the major themes that dominated participants' answers. As a result, we generated a coding tree (Figure 2). According to the coding template, we assigned statements to the themes. A total of 3 of the study authors, all anesthesiology residents GG, GS, and SA, rated the interview statements separately from each other using the coding tree (Figure 2). If the 3 investigators disagreed after multiple data coding, the final decision was taken in a joint discussion. Interrater reliability was calculated to investigate the consistency of the coding tree's application.

XSL•FO
RenderX

Table 1. The most commonly used positive and negative terms to describe Visual Clot.

Positive terms		Negative terms	
Terms	Frequency, n	Terms	Frequency, n
Easy or easier	28	Missing	13
Good	16	Quantitative or quantification	11
Fast or faster	13	Information	9
See a problem	12	Values	8
Interpret quickly	11	Time	8
Simple	9	Fibrinogen	6
Understand	9	Numbers	6
Visual	8	Hyperfibrinolysis	6
Interpretation	7	Less	5
Intuitive	7	Confusing	4
At a glance	7	Flashing	4
Overview	7	Simplified	3

Figure 2. A coding tree representing the themes describing positive and negative user perceptions. ROTEM: rotational thromboelastometry.



Part II: Survey

The literature states that quantitative data can help generalize and confirm specific observations found in qualitative research [29-32]. For the subsequent survey, we defined 5 statements based on the previously identified themes. The same group of interviewed anesthesiologists was asked to rate them on 5-point Likert scales in a questionnaire created using Google Forms (Alphabet Inc). Participants were informed that the survey takes only a few minutes to complete, participation is voluntary, and no compensation is offered. The translated announcement of the survey invitation is displayed in Multimedia Appendix 3. The data collection was finished 3 weeks after the questionnaire was sent.

Statistical Analysis

The interview data analysis and figures were made using Microsoft Word and Excel (Microsoft Corp). We present the number of statements and their percentage distribution in the identified themes.

To define the interrater reliability of the coding template, we calculated Fleiss' Kappa using R (version 4.0.5; R Foundation for Statistical Computing). We calculated every statement's median and IQR for the survey analysis. We used the Wilcoxon signed rank test to determine the difference between the median and neutral answers. Statistical significance was indicated as $P<.05$.

XSL•FO
RenderX

Results

Study and Participant Characteristics

Detailed information on the study and participants is provided in Table 2. Residents and nurses were the dominant participants in the interviews. The most experienced participant had 33 years of experience in anesthesia. The least experienced had less than 1 year. Residents and nurses also dominated the survey.

Table 2. Study and participant characteristics.

Characteristics	Values
Study characteristics	
Total number of interviewed participants, n	77
Total number of participants completed the survey, n	23
Participant characteristics	
Interview participants	
Staff physicians, n (%)	8 (10.4)
Residents, n (%)	35 (45.5)
Anesthesiology nurses, n (%)	34 (44.2)
Anesthesia experience in years, median (IQR)	8 (3-10)
Number of ROTEM[a] interpretations per year, median (IQR)	26 (5-41)
Survey participants	
Staff physicians, n (%)	7 (30.4)
Residents, n (%)	8 (34.8)
Anesthesiology nurses, n (%)	8 (34.8)

[a]ROTEM: rotational thromboelastometry.

Part I: Qualitative Analysis of Interview Answers

Word Count Analysis

The most frequently used words and word combinations used to describe the advantages of Visual Clot were: easier or easy (26/77, 33.8%), interpret or interpretation (23.4%, 18/77 participants), quick or quickly (19.5%, 15/77 participants), visual, visualize, visualized, or visualization (19.5%, 15/77 participants), good (16.9%, 13/77 participants), faster or fast (15.6%, 12/77 participants). In the group of statements describing the limitations of Visual Clot, the words and word groups most frequently used were: ROTEM (23.4%, 18/77 participants), missing (information or values or numbers; 16.9%, 13/77 participants), quantitative or quantification (16.9%, 13/77 participants). Table 1 visually represents the most commonly used words in positive and negative perceptions.

Coding Tree

Figure 2 shows the generated coding tree, including 2 main domains and 9 themes. The interrater reliability of the tree raters was 0.856 (95% CI 0.831-0.880), indicating almost perfect agreement [33].

Statements Describing Visual Clot

Table 3 demonstrates examples of statements assigned to particular subtopics with participant counts and percentages.

A total of 4 comments were defined as positive but not assigned to any themes. There was 1 such statement in the negative group. A total of 19 comments were not assigned to any theme and were described as noncodable.

XSL•FO
RenderX

Table 3. Statements examples assigned to particular domain and subtopics with participant count and percentages.

Major domain and subtopics	Examples
Positive statements describing Visual Clot (179/319, 56.1%)	
Positive design features (63/295, 21.4%)	• "Very simplified" (Participant 13). • "Visual presentation" (Participant 21). • "Tells figuratively what to do" (Participant 26).
Positive usability features (30/295, 10.2%)	• "It can be perfectly integrated in the clinic" (Participant 24). • "A good tool to get an overview" (Participant 32). • "You can see at-a-glance what is missing" (Participant 70).
Intuitive and easy to learn (61/295, 20.7%)	• "Directly applicable and does not require long training" (Participant 5). • "Very intuitive" (Participant 75). • "Easy to understand" (Participant 76).
Time saving (21/295, 7.1%)	• "2-3 seconds to a quick overview" (Participant 22). • "Immediate detection of the problem" (Participant 5). • "Quick to interpret" (Participant 27).
Negative statements describing Visual Clot (113/319, 35.4%)	
Design problems (23/295, 7.8%)	• "Confusing, blinking" (Participant 57). • "Quality of hyperfibrinolysis difficult to demonstrate" (Participant 22). • "Not everything you see is relevant" (Participant 24).
Usability problems (12/295, 4.1%)	• "Quantity of change is not visible" (Participant 2). • "You have to know the pictures first" (Participant 30). • "No prioritization possible" (Participant 76).
Incompleteness: a lack of a quantitative component (62/295, 21.0%)	• "Also values that are in normal range—is it close to the limit or not?" (Participant 54). • "No graduation 'all or nothing'" (Participant 60). • "Exact quantification not possible" (Participant 4).
Need for training (15/295, 5.1%)	• "Not yet established" (Participant 29). • "Needs to get used to it" (Participant 45). • "Needs habituation (not used by default yet)" (Participant 50).
Combination: Visual Clot and ROTEM[a] (8/295, 2.7%)	
	• "Would be nice to have it together with ROTEM" (Participant 50). • "Ideal solution if could be combined with ROTEM" (Participant 49). • "Combination of both necessary" (Participant 74).

[a]ROTEM: rotational thromboelastometry.

Positive Statements Describing Visual Clot

Positive Design Features

Most comments were made on this topic, emphasizing that a "pictorially summarized" (participant 2) and "visually appealing" (participant 3) data presentation allows one to see "the relevant ROTEM information at-a-glance" (participant 9). Such a design supports health care professionals in making clinical decisions. It is essential in emergencies because the actual coagulation status is immediately visible (participant 14) and it is instantly apparent which hemostasis components are missing (participant 12).

Positive Usability Features

Visual Clot is "a good tool for broad application," stated participant 72. It enables "pre-interpretation of the complex information" (participant 3) and focuses "on the essential" (participant 65). The benefits of the Visual Clot in urgent situations were also highlighted: the technology is "very good for emergencies," stated participant 13.

Intuitive and Easy to Learn

As in the previous study [24], the Visual Clot was also described here as intuitive and easy to learn. "Very intuitive, short time needed to understand it,"—pointed out participant 6. It was underlined that visualizations provided by the Visual Clot are "quickly recognizable even by untrained persons or with little knowledge of coagulation" (participant 13).

Time Saving

The Visual Clot provides an "overview at-a-glance," as participant 76 said. "I immediately saw what was missing," stated participant 16. These features lead to quicker diagnosis—"focus is faster on the problem"—as participant 46 said, and thus to faster initiation of treatment.

XSL•FO
RenderX

Negative Statements Describing Visual Clot

Design Problems

Several ideas that could potentially enhance Visual Clot's design were identified. Participant 47 pointed out that the presentation of platelets and fibrinogen are similar, and thus it is difficult to distinguish. Participant 48 also agreed: "I did not notice that platelets were missing because it was white and dashed like fibrinogen." Some participants found that the Visual Clot is too dynamic—too much movement on the screen, which can lead to distraction and make the interpretation of the results difficult "Even if coagulation status is fine, everything is moving, and you can poorly differentiate what is missing" (participant 53), "moves too much, even if everything is fine—distraction" (participant 55).

Usability Problems

Visual Clot is "confusing at the beginning"—stated participant 31 and added that it is "difficult to use without routine." Visual Clot provides "too much information at once,"—participant 53 pointed out.

Incompleteness: Lack of a Quantitative Component

The central Visual Clot aspect criticized was the technology's incompleteness in terms of lacking a quantitative component. Several participants stated that the Visual Clot is "not precise" (participant 1), which can be explained in the words of participants 53 and 9, respectively, who said that in the Visual Clot "quantitative is missing" and that one "can get more information with the ROTEM."

Need for Training

The main point identified in the participants' opinions on this topic was the lack of experience working with this technology and that it is a very new tool not yet established in clinical practice.

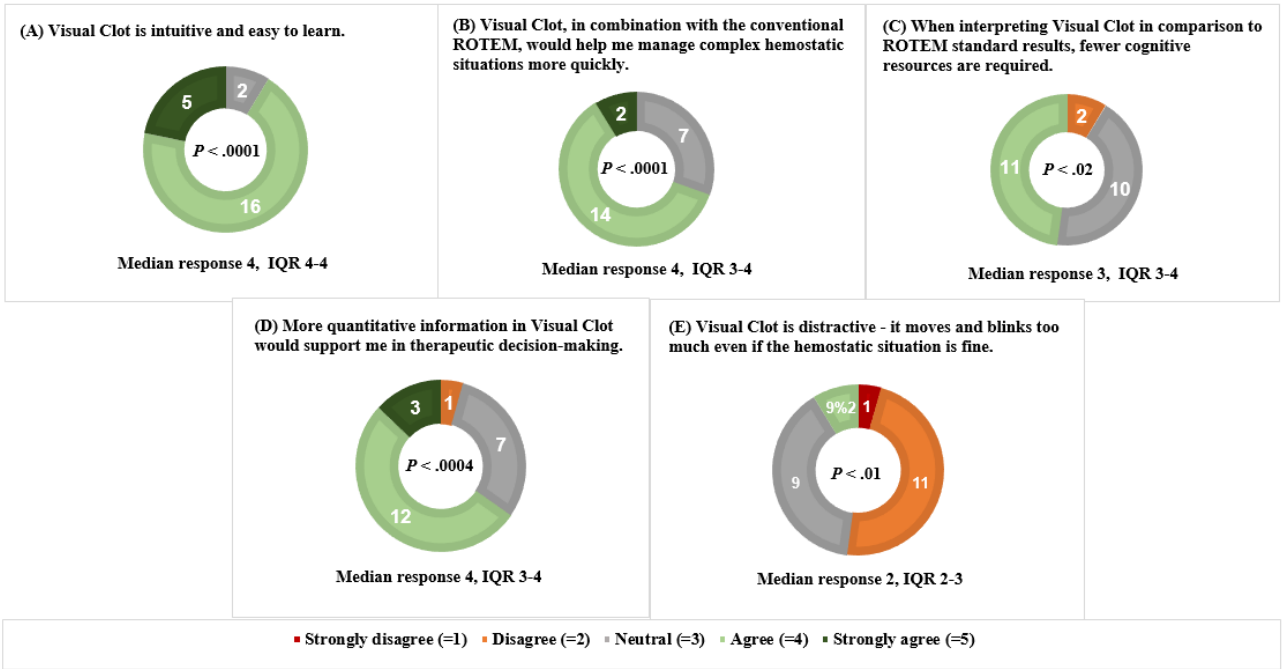
Combination of Visual Clot and ROTEM

Several participants said they could benefit from combining the Visual Clot and ROTEM when interpreting coagulation assays. "A combination of Visual Clot and ROTEM would be perfect," pointed out participant 19, while participant 74 said, "a combination of both is necessary." There was no difference in positive and negative statements based on participants' specialty or level of experience.

Part II: Analysis of Statements Assessed in the Survey

Figure 3 shows the detailed evaluation of the statements rated in the survey.

Figure 3. Pie charts presenting survey results with the number of participants who chose a particular category (N=23). ROTEM: rotational thromboelastometry.



All sample medians differed statistically significantly from neutral ($P<.05$). The number of participants in the quantitative part of the evaluation differs from the qualitative part because not all participants completed the questionnaire. The results are presented as medians and IQR. *P* values are provided to indicate a statistically significant difference between the median of the sample and the neutral value.

Discussion

Principal Findings

This mixed qualitative-quantitative study analyzed the perceptions of anesthesiology personnel regarding Visual Clot—a new situation-awareness and user-oriented visualization technology for viscoelastic hemostatic resuscitation—after the high-fidelity simulation study. User perceptions enable us to identify the positive aspects of the technology and reveal the potential for improvement in the future. After computer-based

XSL•FO
RenderX

studies, this is the first time that Visual Clot has been evaluated in a high-fidelity simulation study, a validated process for testing a noncertified product in an environment that closely resembles clinical reality [34,35].

The principal findings demonstrate that the design features of Visual Clot have received the most positive comments. As in the previous computer-based Visual Clot study [24], the participants of this high-fidelity simulation study emphasized that the way this technology is designed provides a good overview of the clotting situation and is an additional help in the decision-making process during acute bleeding situations. Further, Visual Clot was described as intuitive and easy to learn. Participants repeatedly mentioned that the results of Visual Clot are quickly recognizable and understandable even by inexperienced clinicians. The main criticism concerned the lack of quantitative information.

Previous Visual Clot studies [22-24,36,37] underline the benefits of additional visualization technology, simplifying standard ROTEM data interpretation [20]. Anesthesia providers using Visual Clot in a high-fidelity simulation study were more likely to correctly administer targeted coagulation therapy and to give the first targeted coagulation product faster. In addition, participants demonstrated greater decision-making confidence with Visual Clot [23]. Moreover, the correctness of the clinical decisions was independent of previous rotational thromboelastometry knowledge and experience.

The superior participants' performance when working with Visual Clot may be explained by its design supporting the strengths of human sensory perception. The Visual Clot was developed to assist care providers in managing highly complex coagulation situations, presenting the data in an awareness-oriented interface design. The main aim of this design is to convey the information as quickly as possible and with the lowest cognitive effort [25].

Principles of situation awareness-oriented and user-centered design enables effective data management and a comprehensive understanding of what is happening and thus help to stay situationally aware. This concept is essential in many domains, including medicine, where managing complex and dynamic situations is fundamental [38,39].

Its definition breaks down into three separate phases (1) the perception of environmental elements in the current situation within a volume of time and space, (2) understanding their meaning, and (3) their projection in the near future. Based on this, the Visual Clot data are visually represented, preprocessed, and simplified. The results of coagulation parameters are divided into 3 categories: too low, normal, or too high. Such information presentation increases diagnostic confidence, but numeric indicators are needed for precise data analysis and targeted treatment initiation. As previously indicated, the lack of quantitative information is reflected in user responses. It also explains the participants' considerations that combining Visual Clot and ROTEM would be helpful in clinical decision-making.

Some other technologies based on situation awareness and user-centered design principles include Philips Visual Patient Avatar (Philips) [40], AlertWatch (AlertWatch Inc), Dynamic Lung Panel and PulmoSight (Hamilton Medical AG), HemoSight and Physiology Screen (Mindray Medical International Limited), and Alarm Status Visualizer (Masimo Corp) [41,42].

This study showed user perceptions regarding the new situation awareness-based, user-oriented technology for thromboelastometry data presentation—Visual Clot. It makes us aware of the user's needs and could help us simplify information processing and decision-making in the future. An integral facet of advancing the technological framework informed by the results of this study lies in the prospect of merging quantitative data into the Visual Clot platform and presenting this merged information in a consolidated interface. This concerted integration promises to align both quantitative and qualitative data to provide a more complete and accurate representation of prevailing conditions. This integration can be achieved in a variety of ways, including the direct overlay of numerical values onto the Visual Clot visualization, or the parallel juxtaposition of a complementary graphical representation alongside the numerical data set.

Strengths and Limitations

This study has several strengths and limitations. The interview part of the study has the typical limitations of qualitative research. The findings of qualitative analysis cannot be extrapolated to larger populations with the same certainty as quantitative results because the findings are on the subjective basis and not tested for statistical significance [43]. However, the quantitative survey helped to provide greater insight into the importance of the main themes identified. Moreover, the interviewed participants were selected according to their availability in the clinical praxis and not randomly.

Furthermore, the number of participants in the survey was lower than in the interviews because not all participants in the simulation study completed the survey. Finally, it is a single-center study performed in a university hospital with high care standards in Europe. User perceptions may vary across diverse clinical settings in different parts of the world.

Conclusions

After previous studies investigating user perceptions of Visual Clot in computer-based simulation studies, this is the first study to analyze the user perceptions of Visual Clot in a high-fidelity simulation—the intermediate step between computer-based simulation studies in a laboratory and real-life use. In this study, Visual Clot appeared to be a well-accepted additional tool supporting health care professionals working with ROTEM. Based on participants' perceptions, user-centered and situation awareness-oriented design, as shown in Visual Clot, can simplify the presentation of complex information and thus make critical decision-making quicker and more efficient. The benefits of this technology have been particularly highlighted in emergencies and even for care providers with little experience in coagulation management. Participants described Visual Clot as intuitive and easy to learn. The lack of a quantitative component has been identified as a significant limitation. These findings highlight the advantages of Visual Clot and its potential

XSL•FO
RenderX

for improvement may help further develop this and other situation awareness-based technologies.

Acknowledgments

The authors are thankful to the study participants for their time and effort. The Institute of Anesthesiology of the University Hospital of Zurich, Zurich, Switzerland and the University of Zurich, Zurich, Switzerland funded this study.

Data Availability

The data sets used and analyzed during this study are available from the corresponding author on reasonable request.

Authors' Contributions

GG, AM, CC, ADB, GS, MK, BG, CBN, TRR, DWT, and SA contributed to the conceptualization; GG, AM, CC, ADB, GS, MK, BG, CBN, TRR, DWT, and SA contributed to the methodology; GG, GS, and SA performed the formal analysis; AM, CC, ADB, GS, TRR, and DWT contributed to the investigation; GG, GS, and SA performed data curation; GG and DWT contributed to writing—original draft preparation; GG, DRS, CBN, DWT, SA, and DF contributed to writing—review and editing; GG and DWT contributed to visualization; DWT and SA performed supervision; and DWT performed project administration.

Conflicts of Interest

DRS's academic department is receiving grant support from the Swiss National Science Foundation, Berne, Switzerland, the Swiss Society of Anesthesiology and Perioperative Medicine, Berne, Switzerland, the Swiss Foundation for Anesthesia Research, Zurich, Switzerland, Vifor SA, Villars-sur-Glâne, Switzerland and Vifor (International) AG, St. Gallen, Switzerland. DRS is cochair of the ABC-Trauma Faculty, sponsored by unrestricted educational grants from Novo Nordisk Health Care AG, Zurich, Switzerland, CSL Behring GmbH, Marburg, Germany, LFB Biomédicaments, Courtaboeuf Cedex, France and Octapharma AG, Lachen, Switzerland. DRS received honoraria or travel support for consulting or lecturing from Alliance Rouge, Bern, Switzerland, Danube University of Krems, Austria, European Society of Anesthesiology and Intensive Care, Brussels, BE, International Foundation for Patient Blood Management, Basel, Switzerland, Korean Society of Anesthesiologists, Seoul, Korea, Network for the Advancement of Patient Blood Management, Haemostasis and Thrombosis, Paris, France, Society for the Advancement of Blood Management, Mount Royal NJ, Alexion Pharmaceuticals Inc, Boston, MA, AstraZeneca AG, Baar, Switzerland, Bayer AG, Zürich, Switzerland, B. Braun Melsungen AG, Melsungen, Germany, Baxter AG, Glattpark, Switzerland, CSL Behring GmbH, Hattersheim am Main, Germany and Berne, Switzerland, CSL Vifor (Switzerland) Villars-sur-Glâne, Switzerland, CSL Vifor (International), St Gallen, Switzerland, Celgene International II Sàrl, Couvet, Switzerland, Daiichi Sankyo AG, Thalwil, Switzerland, Haemonetics, Braintree, Massachusetts, United States, LFB Biomédicaments, Courtaboeuf Cedex, France, Merck Sharp & Dohme, Kenilworth, New Jersey, United States, Novo Nordisk Health Care AG, Zurich, Switzerland, Octapharma AG, Lachen, Switzerland, Pharmacosmos A/S, Holbaek, Denmark, Pierre Fabre Pharma, Alschwil, Switzerland, Portola Schweiz GmbH, Aarau, Switzerland, Roche Diagnostics International Ltd, Reinach, Switzerland, Sarstedt AG & Co, Sevelen, Switzerland and Nümbrecht, Germany, Shire Switzerland GmbH, Zug, Switzerland, Takeda, Glattpark, Switzerland, Werfen, Bedford, MA, Zuellig Pharma Holdings, Singapore, Singapore. CBN is an inventor of Visual Patient and Visual Patient Predictive technologies, for which the University of Zurich and Koninklijke Philips N.V. hold patents, patent applications, design protections, and trademarks. Joint-development and licensing agreements exist with Philips Medizin Systeme Böblingen GmbH, Böblingen, Germany; Koninklijke Philips N.V., Amsterdam, The Netherlands; Philips Research/Philips Electronics Nederland BV, Eindhoven, The Netherlands; and Philips USA, Cambridge, Massachusetts, United States. Within the framework of these agreements, CBN receives travel support, lecturing and consulting honoraria, and may potentially receive royalties in the event of successful commercialization. CBN is an inventor of Visual Clot technology, with patent applications, design protections, and trademarks held by the University of Zurich. In case of successful commercialization, CBN may receive royalties. CBN is an inventor of Visual Blood technology, for which the University of Zurich holds patent applications and design protections; potential royalties may follow successful commercialization. CBN received travel support, lecturing, and consulting honoraria from Instrumentation Laboratory—Werfen, Bedford, Massachusetts, United States. DWT is the first named inventor of Visual Patient and Visual Patient Predictive technologies, for which the University of Zurich and Koninklijke Philips N.V. hold patents, patent applications, design protections, and trademarks. Joint-development and licensing agreements exist with Philips Medizin Systeme Böblingen GmbH, Böblingen, Germany; Koninklijke Philips N.V., Amsterdam, The Netherlands; Philips Research/Philips Electronics Nederland BV, Eindhoven, The Netherlands; and Philips USA, Cambridge, Massachusetts, United States. Within the framework of these agreements, DWT receives research funding, travel support, lecturing and consulting honoraria, and may potentially receive royalties in the event of successful commercialization. DWT also holds a position on the Philips Patient Safety Advisory Board. DWT is the first named inventor of Visual Clot technology, with patent applications, design protections, and trademarks held by the University of Zurich. In case of successful commercialization, DWT may receive royalties. DWT is the first named inventor of Visual Blood technology, for which the University of Zurich holds patent applications and design protections; potential royalties may follow successful commercialization. Additionally, DWT received travel support, lecturing, and consulting honoraria

XSL•FO
RenderX

from Instrumentation Laboratory—Werfen, Bedford, Massachusetts, United States, the Swiss Foundation for Anaesthesia Research in Zurich, Switzerland, and the International Symposium on Intensive Care and Emergency Medicine in Brussels, Belgium. No other funding or competing interests declared.

Multimedia Appendix 1
Educational Visual Clot video.
[MOV File , 19983 KB - humanfactors_v11i1e47991_app1.mov]

Multimedia Appendix 2
Complete translated field notes of participant interviews.
[PDF File (Adobe PDF File), 669 KB - humanfactors_v11i1e47991_app2.pdf]

Multimedia Appendix 3
Translated announcement of the survey invitation.
[PDF File (Adobe PDF File), 204 KB - humanfactors_v11i1e47991_app3.pdf]

References

1. Spahn DR, Bouillon B, Cerny V, Duranteau J, Filipescu D, Hunt BJ, et al. The European guideline on management of major bleeding and coagulopathy following trauma: fifth edition. Crit Care 2019;23(1):98 [FREE Full text] [doi: 10.1186/s13054-019-2347-3] [Medline: 30917843]
2. Weber CF, Zacharowski K. [Point of care 2.0: coagulation monitoring using Rotem® Sigma and Teg® 6S]. Anasthesiol Intensivmed Notfallmed Schmerzther 2018;53(6):412-424. [doi: 10.1055/s-0043-107755] [Medline: 29945283]
3. Riehl K, Lefering R, Maegele M, Caspers M, Migliorini F, Schenker H, et al. Is ROTEM diagnostic in trauma care associated with lower mortality rates in bleeding patients?-A retrospective analysis of 7461 patients derived from the TraumaRegister DGU®. J Clin Med 2022;11(20):6150 [FREE Full text] [doi: 10.3390/jcm11206150] [Medline: 36294471]
4. American Society of Anesthesiologists Task Force on Perioperative Blood Management. Practice guidelines for perioperative blood management: an updated report by the American Society of Anesthesiologists Task Force on perioperative blood management*. Anesthesiology 2015;122(2):241-275 [FREE Full text] [doi: 10.1097/ALN.0000000000000463] [Medline: 25545654]
5. Tibi P, McClure RS, Huang J, Baker RA, Fitzgerald D, Mazer CD, et al. STS/SCA/AmSECT/SABM update to the clinical practice guidelines on patient blood management. Ann Thorac Surg 2021;112(3):981-1004 [FREE Full text] [doi: 10.1016/j.athoracsur.2021.03.033] [Medline: 34217505]
6. Collins PW, Lilley G, Bruynseels D, Laurent DBS, Cannings-John R, Precious E, et al. Fibrin-based clot formation as an early and rapid biomarker for progression of postpartum hemorrhage: a prospective study. Blood 2014;124(11):1727-1736 [FREE Full text] [doi: 10.1182/blood-2014-04-567891] [Medline: 25024304]
7. McNamara H, Kenyon C, Smith R, Mallaiah S, Barclay P. Four years' experience of a ROTEM® -guided algorithm for treatment of coagulopathy in obstetric haemorrhage. Anaesthesia 2019;74(8):984-991 [FREE Full text] [doi: 10.1111/anae.14628] [Medline: 30950521]
8. Wikkelsø A, Wetterslev J, Møller AM, Afshari A. Thromboelastography (TEG) or thromboelastometry (ROTEM) to monitor haemostatic treatment versus usual care in adults or children with bleeding. Cochrane Database Syst Rev 2016;2016(8):CD007871 [FREE Full text] [doi: 10.1002/14651858.CD007871.pub3] [Medline: 27552162]
9. Hawkins RB, Raymond SL, Hartjes T, Efron PA, Larson SD, Andreoni KA, et al. Review: the perioperative use of thromboelastography for liver transplant patients. Transplant Proc 2018;50(10):3552-3558. [doi: 10.1016/j.transproceed.2018.07.032] [Medline: 30577236]
10. Weber CF, Görlinger K, Meininger D, Herrmann E, Bingold T, Moritz A, et al. Point-of-care testing: a prospective, randomized clinical trial of efficacy in coagulopathic cardiac surgery patients. Anesthesiology 2012;117(3):531-547 [FREE Full text] [doi: 10.1097/ALN.0b013e318264c644] [Medline: 22914710]
11. Görlinger K, Dirkmann D, Hanke AA, Kamler M, Kottenberg E, Thielmann M, et al. First-line therapy with coagulation factor concentrates combined with point-of-care coagulation testing is associated with decreased allogeneic blood transfusion in cardiovascular surgery: a retrospective, single-center cohort study. Anesthesiology 2011;115(6):1179-1191 [FREE Full text] [doi: 10.1097/ALN.0b013e31823497dd] [Medline: 21970887]
12. Kvint S, Schuster J, Kumar MA. Neurosurgical applications of viscoelastic hemostatic assays. Neurosurg Focus 2017;43(5):E9 [FREE Full text] [doi: 10.3171/2017.8.FOCUS17447] [Medline: 29088950]
13. Welling H, Ostrowski SR, Stensballe J, Vestergaard MR, Partoft S, White J, et al. Management of bleeding in major burn surgery. Burns 2019;45(4):755-762. [doi: 10.1016/j.burns.2018.08.024] [Medline: 30292526]
14. Sahli SD, Pedrazzi N, Braun J, Spahn DR, Kaserer A, Plock JA. Effect of a factor-based coagulation management on blood product use after major burn injury: a retrospective cohort study. Burns 2021;47(7):1486-1494 [FREE Full text] [doi: 10.1016/j.burns.2021.08.015] [Medline: 34465509]

XSL•FO
RenderX

15. Treliński J, Misiewicz M, Robak M, Smolewski P, Chojnowski K. Assessment of rotation thromboelastometry (ROTEM) parameters in patients with multiple myeloma at diagnosis. Thromb Res 2014;133(4):667-670. [doi: 10.1016/j.thromres.2014.01.011] [Medline: 24451990]
16. Young GA, Carmona R, Garcia VC. Thromboelastography and thrombin generation assay in inherited afibrinogenemia. Haemophilia 2018;24(6):e410-e416. [doi: 10.1111/hae.13620] [Medline: 30345592]
17. Chitlur M, Rivard GE, Lillicrap D, Mann K, Shima M, Young G, Factor VIII, Factor IX, and Rare Coagulation Disorders Subcommittee of the Scientific and Standardisation Committee of the International Society on Thrombosis and Haemostasis. Recommendations for performing thromboelastography/thromboelastometry in hemophilia: communication from the SSC of the ISTH. J Thromb Haemost 2014;12(1):103-106 [FREE Full text] [doi: 10.1111/jth.12458] [Medline: 24261669]
18. Young G, Sørensen B, Dargaud Y, Negrier C, Brummel-Ziedins K, Key NS. Thrombin generation and whole blood viscoelastic assays in the management of hemophilia: current state of art and future perspectives. Blood 2013;121(11):1944-1950 [FREE Full text] [doi: 10.1182/blood-2012-08-378935] [Medline: 23319573]
19. Srivastava A, Kelleher A. Point-of-care coagulation testing. Contin Educ Anaesth Crit Care Pain 2013;13(1):12-16 [FREE Full text] [doi: 10.1093/bjaceaccp/mks049]
20. Gasciauskaite G, Malorgio A, Castellucci C, Budowski A, Schweiger G, Kolbe M, et al. User perceptions of ROTEM-guided haemostatic resuscitation: a mixed qualitative-quantitative study. Bioengineering (Basel) 2023;10(3):386 [FREE Full text] [doi: 10.3390/bioengineering10030386] [Medline: 36978777]
21. Dixon A, Underwood S, Schreiber M. Implementing thrombelastography: experiences from a level I trauma institution. Transfusion 2020;60(Suppl 6):S29-S32 [FREE Full text] [doi: 10.1111/trf.16070] [Medline: 33089931]
22. Rössler J, Meybohm P, Spahn DR, Zacharowski K, Braun J, Nöthiger CB, et al. Improving decision making through presentation of viscoelastic tests as a 3D animated blood clot: the visual clot. Anaesthesia 2020;75(8):1059-1069 [FREE Full text] [doi: 10.1111/anae.14985] [Medline: 32030729]
23. Castellucci C, Malorgio A, Budowski AD, Akbas S, Kolbe M, Grande B, et al. Coagulation management of critically bleeding patients with viscoelastic testing presented as a 3D-animated blood clot (the visual clot): randomized controlled high-fidelity simulation study. J Med Internet Res 2023;25:e43895 [FREE Full text] [doi: 10.2196/43895] [Medline: 37824182]
24. Roche TR, Said S, Rössler J, Gozdzik M, Meybohm P, Zacharowski K, et al. Physicians' perceptions of a situation awareness-oriented visualization technology for viscoelastic blood coagulation management (visual clot): mixed methods study. JMIR Serious Games 2020;8(4):e19036 [FREE Full text] [doi: 10.2196/19036] [Medline: 33172834]
25. Endsley MR. Designing for Situation Awareness: An Approach to User-Centered Design, Second Edition. Boca Raton, FL, USA: CRC Press Inc; 2011.
26. Wittgenstein L. Tractatus Logico-Philosophicus. London, UK: Routledge & Kegan Paul; 1922.
27. Degani A, Jorgensen C, Iverson D, Shafto M, Olson L. On Organization of Information: Approach and Early Work. Moffett Field, California: National Aeronautics and Space Administration, Ames Research Center; 2009.
28. Symon G, Cassell C. Qualitative Organizational Research: Core Methods and Current Challenges. Thousand Oaks, CA: SAGE Publications; 2012.
29. Guetterman TC, Fetters MD, Creswell JW. Integrating quantitative and qualitative results in health science mixed methods research through joint displays. Ann Fam Med 2015;13(6):554-561 [FREE Full text] [doi: 10.1370/afm.1865] [Medline: 26553895]
30. Smajic E, Avdic D, Pasic A, Prcic A, Stancic M. Mixed methodology of scientific research in healthcare. Acta Inform Med 2022;30(1):57-60 [FREE Full text] [doi: 10.5455/aim.2022.30.57-60] [Medline: 35800901]
31. Kaur N, Vedel I, El Sherif R, Pluye P. Practical mixed methods strategies used to integrate qualitative and quantitative methods in community-based primary health care research. Fam Pract 2019;36(5):666-671 [FREE Full text] [doi: 10.1093/fampra/cmz010] [Medline: 30888015]
32. Webster CS, McKillop A, Bennett W, Bagg W. A qualitative and semiquantitative exploration of the experience of a rural and regional clinical placement programme. Med Sci Educ 2020;30(2):783-789 [FREE Full text] [doi: 10.1007/s40670-020-00949-6] [Medline: 34457735]
33. Landis JR, Koch GG. The measurement of observer agreement for categorical data. Biometrics 1977;33(1):159-174. [Medline: 843571]
34. Merry AF, Hannam JA, Webster CS, Edwards KE, Torrie J, Frampton C, et al. Retesting the hypothesis of a clinical randomized controlled trial in a simulation environment to validate anesthesia simulation in error research (the VASER study). Anesthesiology 2017;126(3):472-481 [FREE Full text] [doi: 10.1097/ALN.0000000000001514] [Medline: 28079566]
35. Kurup V, Matei V, Ray J. Role of in-situ simulation for training in healthcare: opportunities and challenges. Curr Opin Anaesthesiol 2017;30(6):755-760. [doi: 10.1097/ACO.0000000000000514] [Medline: 28968283]
36. Castellucci C, Braun J, Said S, Roche TR, Nöthiger CB, Spahn DR, et al. Faster time to treatment decision of viscoelastic coagulation test results through improved perception with the animated visual clot: a multicenter comparative eye-tracking study. Diagnostics (Basel) 2022;12(5):1269 [FREE Full text] [doi: 10.3390/diagnostics12051269] [Medline: 35626425]

XSL•FO
RenderX

37. Said S, Roche TR, Braun J, Ganter MT, Meybohm P, Herrmann J, et al. Effects of an animated blood clot technology (visual clot) on the decision-making of users inexperienced in viscoelastic testing: multicenter trial. J Med Internet Res 2021;23(5):e27124 [FREE Full text] [doi: 10.2196/27124] [Medline: 33843602]
38. Schulz C, Endsley MR, Kochs EF, Gelb AW, Wagner KJ. Situation awareness in anesthesia: concept and research. Anesthesiology 2013;118(3):729-742 [FREE Full text] [doi: 10.1097/ALN.0b013e318280a40f] [Medline: 23291626]
39. Kang J, Hu J, Yan C, Xing X, Tu S, Zhou F. Development and applications of the anaesthetists' non-technical skills behavioural marker system: protocol for a systematic review. BMJ Open 2022;12(12):e065519 [FREE Full text] [doi: 10.1136/bmjopen-2022-065519] [Medline: 36517093]
40. Patient monitoring: visual patient avatar. Philips. 2023. URL: https://www.philips.co.uk/healthcare/technology/visual-patient-avatar [accessed 2023-12-22]
41. Products. mindray. 2023. URL: https://www.mindray.com/uk/products [accessed 2023-12-22]
42. Gasciauskaite G, Lunkiewicz J, Roche TR, Spahn DR, Nöthiger CB, Tscholl DW. Human-centered visualization technologies for patient monitoring are the future: a narrative review. Crit Care 2023;27(1):254 [FREE Full text] [doi: 10.1186/s13054-023-04544-0] [Medline: 37381008]
43. Atieno OA. An analysis of the strengths and limitation of qualitative and quantitative research paradigms. Probl Educ 21st Century 2009;13:13-38 [FREE Full text]

Abbreviations

NASA: National Aeronautics and Space Administration
ROTEM: rotational thromboelastometry

Edited by A Kushniruk; submitted 07.04.23; peer-reviewed by D Smriti, M Endsley; comments to author 23.09.23; revised version received 17.10.23; accepted 27.11.23; published 11.01.24.

Please cite as:
Gasciauskaite G, Castellucci C, Malorgio A, Budowski AD, Schweiger G, Kolbe M, Fries D, Grande B, Nöthiger CB, Spahn DR, Roche TR, Tscholl DW, Akbas S
User Perceptions of Visual Clot in a High-Fidelity Simulation Study: Mixed Qualitative-Quantitative Study
JMIR Hum Factors 2024;11:e47991
URL: https://humanfactors.jmir.org/2024/1/e47991
doi:10.2196/47991
PMID:38206666

©Greta Gasciauskaite, Clara Castellucci, Amos Malorgio, Alexandra D Budowski, Giovanna Schweiger, Michaela Kolbe, Daniel Fries, Bastian Grande, Christoph B Nöthiger, Donat R Spahn, Tadzio R Roche, David W Tscholl, Samira Akbas. Originally published in JMIR Human Factors (https://humanfactors.jmir.org), 11.01.2024. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on https://humanfactors.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Digital Care Pathway for Patients With Sleep Apnea in Specialized Care: Mixed Methods Study

Jari Haverinen[1,2], MSc, MHS; Terttu Harju[3], MD, PhD; Hanna Mikkonen[4], RNC; Pia Liljamo[5], RNC, PhD; Miia Turpeinen[6], MD, PhD; Jarmo Reponen[2,6], MD, PhD

[1]Finnish Coordinating Center for Health Technology Assessment, Oulu University Hospital, Oulu, Finland
[2]FinnTelemedicum, Research Unit of Health Sciences and Technology, Faculty of Medicine, University of Oulu, Oulu, Finland
[3]Medical Research Center Oulu, Oulu Pulmonary Department, Oulu University Hospital and University of Oulu, Oulu, Finland
[4]The Wellbeing Services County of North Ostrobothnia, Oulu, Finland
[5]Finnish Institute for Health and Welfare, Department of Knowledge Brokers, Data and Analytics Unit, Helsinki, Finland
[6]Medical Research Center Oulu, Oulu University Hospital and University of Oulu, Oulu, Finland

Corresponding Author:
Jari Haverinen, MSc, MHS
Finnish Coordinating Center for Health Technology Assessment
Oulu University Hospital
Kajaanintie 50
Oulu, PL10
Finland
Phone: 358 504095446
Email: jari.haverinen@pohde.fi

Abstract

Background: Sleep apnea is a significant public health disorder in Finland, with a prevalence of 3.7%. Continuous positive airway pressure (CPAP) therapy is the first-line treatment for moderate or severe sleep apnea. From November 18, 2019, all patients who started their CPAP therapy at Oulu University Hospital were attached to a sleep apnea digital care pathway (SA-DCP) and were instructed on its use. Some patients still did not use the SA-DCP although they had started their CPAP therapy.

Objective: We aimed to study health care professionals' (HCPs') perspectives on the SA-DCP and its usefulness for their work; whether the main targets of SA-DCP can be reached: shortening the initial guiding sessions of CPAP therapy, reducing patient calls and contact with HCPs, and improving patients' adherence to CPAP therapy; and patients' perspectives on the SA-DCP and its usefulness to them.

Methods: Overall, 6 HCPs were interviewed in May and June 2021. The survey for SA-DCP users (58/91, 64%) and SA-DCP nonusers (33/91, 36%) was conducted in 2 phases: from May to August 2021 and January to June 2022. CPAP device remote monitoring data were collected from SA-DCP users (80/170, 47.1%) and SA-DCP nonusers (90/170, 52.9%) in May 2021. The registered phone call data were collected during 2019, 2020, and 2021. Feedback on the SA-DCP was collected from 446 patients between February and March 2022.

Results: According to HCPs, introducing the SA-DCP had not yet significantly improved their workload and work practices, but it had brought more flexibility in some communication situations. A larger proportion of SA-DCP users familiarized themselves with prior information about CPAP therapy before the initial guiding session than nonusers (43/58, 74% vs 16/33, 49%; P=.02). Some patients still had not received prior information about CPAP therapy; therefore, most of the sessions were carried out according to their needs. According to the patient survey and remote monitoring data of CPAP devices, adherence to CPAP therapy was high for both SA-DCP users and nonusers. The number of patients' phone calls to HCPs did not decrease during the study. SA-DCP users perceived their abilities to use information and communications technology to be better than nonusers (mean 4.2, SD 0.8 vs mean 3.2, SD 1.2; P<.001).

Conclusions: According to this study, not all the goals set for the introduction of the SA-DCP have been achieved. Despite using the SA-DCP, some patients still wanted to communicate with HCPs by phone. The most significant factors explaining the nonuse of the SA-DCP were lower digital literacy and older age of the patients. In the future, more attention should be paid to these user groups when designing and introducing upcoming digital care pathways.

XSL•FO
RenderX

(JMIR Hum Factors 2024;11:e47809) doi:10.2196/47809

KEYWORDS

health services; telehealth; telemedicine; health personnel; sleep apnea syndromes; mobile phone

Introduction

Background

Sleep apnea is a significant public health disorder in Finland, with a prevalence of 3.7%. The prevalence of sleep apnea worldwide has been increasing in relation to the obesity pandemic [1]. Untreated sleep apnea increases cardiovascular diseases, accidents, likelihood of taking sick leave, and premature mortality [2]. The clinical severity of sleep apnea is defined based on 3 components: daytime sleepiness owing to sleep apnea, the apnea-hypopnea index (AHI), and arterial blood oxygen saturation [2]. Continuous positive airway pressure (CPAP) therapy is the first-line treatment for moderate or severe sleep apnea in addition to conservative therapy (ie, weight loss, avoidance of sleep-disturbing substances, and lifestyle issues) [2]. CPAP therapy is a safe and efficient treatment for sleep apnea, relieving both daytime and nighttime symptoms and improving traffic safety [3,4]. In Finland, the need for CPAP treatment and the number of outpatient visits in both specialized and primary care have increased considerably because of the increased number of patients with sleep apnea [5].

The digitalization of health care has been seen as a potential option for offering treatment to patients regardless of time and place and involving them in their own care [6,7]. In addition, digitalization has the potential to make health care systems more efficient [8]. Despite its potential to improve health care services, digitalization does not automatically guarantee better services [9]. It has been noted as a problem, for example, that digital services are not necessarily aligned with clinician and patient preferences [9]. The challenge is that, in some cases, they complement rather than substitute the current services, and care processes are not always redesigned to achieve the best benefits from digital services [9-12]. Citizens' willingness and ability to use electronic services is also an obstacle to realizing the benefits of health care digitalization [13,14]. Because data breaches cause potentially catastrophic consequences, information security concerns have weakened patients' adoption of digital health services [15,16]. The challenges of the technical implementation of digital services, such as missing functionalities and lack of interoperability with existing information systems, have weakened the willingness of health care professionals (HCPs) to use them [10].

Factors promoting the adoption of digital health services are their perceived benefits for patients and patients' previous positive experiences with electronic services [13,15]. Previous studies showed that digital health interventions can improve patients' adherence to their care [17,18]. For example, Aardoom et al [18] showed that adherence to CPAP therapy in patients with sleep apnea can be improved with digital interventions in the initial months of treatment. Adherence to the use of the digital health service has also been found in some studies to positively affect outcomes [19,20]. Good digital literacy promotes the use of digital health services; studies have found young people have better digital literacy than older age groups [13,21]. As the user base of digital health care services can be very broad, and users can have functional limitations owing to age or illness, the ease of use of these services is important in promoting their use [15,22].

Finland's first phase of health care digitalization involved the digitalizing of HCPs' tools, such as electronic patient records; e-prescribing and digitalization have progressed well [23]. Currently, Finnish citizens are increasingly offered digital health care services and products [22,24]. Several countries, including Finland, have introduced new health technology assessment methods to ensure that digital health provides evidence-based benefits [16,22,25]. Digital care pathways (DCPs) are an example of digital health care services, and today, there are >300 DCPs in use in Finnish specialized care units [26]. One of the main goals of DCPs is to complement or replace traditional health care appointment visits [26]. In addition, DCPs aim to support and help in the self-treatment of long-term illnesses, monitoring, and adaptation to the illness, as well as enable patients to prepare for various health care procedures beforehand [26]. Several DCPs have been studied in Finland from the perspective of HCPs, organizations, and patients [7,10,13,27-31]. One of these DCPs is the sleep apnea DCP (SA-DCP), which was introduced at Oulu University Hospital (OUH) on November 18, 2019 [32]. All patients who start their CPAP therapy in OUH will be attached to the SA-DCP, that is, their patient data will be recorded in it, and they will be instructed on how to log in and use it [32]. When a patient starts on the SA-DCP, they register as an SA-DCP user through strong identification by accepting the terms of use and privacy statement and entering his or her contact information [33].

Objectives

In OUH, the CPAP therapy for patients with sleep apnea begins with an initial guiding session where patients are instructed on using their CPAP device. SA-DCP contains information and instructions about CPAP therapy; therefore, it would be desirable for patients to familiarize themselves with that information in advance [32]. In this way, the initial guiding session of CPAP therapy could be shortened because the basic information about CPAP therapy would not need to be reviewed again during the sessions. The SA-DCP contains reliable information about sleep apnea, its treatment, and CPAP therapy [32]. With the introduction of SA-DCP, it would be desirable to reduce patients' phone calls and other contacts with HCPs when information can be found in the SA-DCP. The SA-DCP also includes electronic messaging between patients and HCPs, which could reduce such calls [32]. The major aim of the SA-DCP is to increase patients' adherence to CPAP therapy. However, there is still a challenge in that some patients with sleep apnea do not log in and use it.

The main aims of the study are as follows:

XSL•FO
RenderX

1. To investigate HCPs' perspectives on the SA-DCP and its usefulness for their work
2. To determine whether the main targets of SA-DCP can be reached: shortening the initial guiding sessions of CPAP therapy, reducing patient calls and contact with HCP, and increasing patients' adherence to CPAP therapy.
3. To examine patients' perspectives on the SA-DCP and its usefulness.

Methods

Study Participants and Data Collection

The study population included HCPs at the OUH and patients who had started their CPAP therapy at the OUH. The patient population consisted of 2 groups. *SA-DCP users* were patients who had registered with the SA-DCP. *SA-DCP nonusers* referred to patients who had not registered with the SA-DCP.

Interviews of HCPs

HCPs of the OUH were contacted via email. Overall, 6 HCPs participated in the interviews from May to June 2021. Of these, 4 (67%) HCPs worked with patients, 1 (17%) was a supervisor, and 1 (17%) connected patients with sleep apnea to the SA-DCP and booked their appointments. The interviews were conducted remotely using a structured questionnaire. The HCPs provided voluntary informed consent for the interview by submitting a signed document. The interviews were then recorded and transcribed.

Survey for Patients With Sleep Apnea

The first part of material collection was conducted between May and August 2021. With the help of OUH HCPs, the survey, along with an invitation to participate and information about it, was sent to SA-DCP nonusers by mail. Respondents could send their responses by prepaid mail or electronically using Webropol Ltd's Webropol survey tool. SA-DPC users were informed about the study through the SA-DCP. They provided their consent and answered the survey using the SA-DCP questionnaire.

The second part of material collection was conducted between January and June 2022. Both SA-DCP users and SA-DCP nonusers were informed about the study with the annual device delivery in an assistive equipment center (AEC). They could send their responses by prepaid mail or answer electronically using Webropol Ltd's Webropol survey tool.

The patients' survey included multiple choice questions, 5-item Likert-type questions (with choices ranging from strongly disagree to strongly agree), and open-ended questions. In total, 33 SA-DCP nonusers and 58 SA-DCP users responded to the survey.

Remote Monitoring Data of CPAP Devices

Information about patients' adherence to CPAP therapy was collected from the remote monitoring data of CPAP devices. The HCP of OUH carried out the material collection manually in May 2021 in connection with 1-year controls of CPAP therapy. The collected information was anonymized and provided to the researchers. In total, CPAP remote monitoring data were collected from 90 SA-DCP nonusers and 80 SA-DCP users.

Registered Data of Phone Calls

The information about the number of patients' phone calls per year to an AEC was collected from Aurora Innovation Ltd's TeleQ program. The registered phone call data were collected during 2019, 2020, and 2021.

SA-DCP Customer Feedback Survey

Patients using the SA-DCP had the opportunity to provide customer feedback using the SA-DCP survey tool. The patients provided informed consent through the SA-DCP that their customer feedback could also be used for research purposes. The customer feedback did not contain any personal information. Feedback on the SA-DCP from 446 patients between February 18 and March 24, 2022, was included in this study.

Statistical Methods

Patients' survey data were analyzed using SPSS software (version 28.0; IBM Corp). Descriptive statistics were applied to calculate the mean and SD for continuous data and frequency and percentage for categorical data. Baseline differences between the groups were explored using a 2-tailed independent sample *t* test for continuous variables and chi-square test for categorical variables. A P value <.05 was considered statistically significant for all analyses.

Qualitative Analysis

Qualitative methods were used in this study to analyze the open-ended questions in patient surveys and interviews with HCPs. The collected material was first analyzed using an inductive content analysis method to obtain a comprehensive understanding [34]. Initially, the HCPs' and patients' responses to the open-ended questions were open coded. Subsequently, the analyzed data were grouped into subcategories, and then similar findings were combined into the main categories to enable the final analysis. Finally, the textual data were analyzed using the quantification method [35].

Ethical Considerations

The study followed the guidelines of the Finnish Advisory Board on Research Integrity [36]. According to Finnish Law (488/1999), this study was exempted from review by the institutional review board (ethics committee of Northern Ostrobothnia Hospital District). The respondents were informed of the study. All participants voluntarily participated in the study and provided their informed consent. The results were processed such that no participants were identifiable in the results or quotations of this study. Sensitive personal information was not collected. The data were processed and stored in a secure environment according to the procedures of the University of Oulu.

Results

The Number of New CPAP Therapies, SA-DCP Users, and Phone Calls in the Years Studied

The number of new CPAP therapies in OUH between 2019 and 2021 is presented in Table 1. The percentage of SA-DCP users

XSL•FO
RenderX

has increased annually, but there are still patients who do not use the SA-DCP (Table 1). The number of phone calls per year to an AEC is presented in Table 1.

Table 1. New continuous positive airway pressure (CPAP) therapies, patients attached to the sleep apnea digital care pathway (SA-DCP), phone calls to an assistive equipment center (AEC) per year, and percentage of SA-DCP users.

	Year		
	2019	2020	2021
The number of new CPAP therapies	1172	1645	1160
The number of patients attached to the SA-DCP	292[a]	1645	1160
The number of SA-DCP users	130	1006	935
Percentage of SA-DCP users	44.5	61.2	80.6
The number of phone calls per year to an AEC	2784	4068	4020

[a]The SA-DCP was introduced from November 18, 2019, onward.

HCPs' Perspectives on the SA-DCP and its Usefulness for Their Work

On the basis of the interviews with HCPs, the main themes, facilitators, and barriers related to using the SA-DCP are presented in Textbox 1. According to the interviewed HCPs, they were unable to identify significant changes in their workload and working practices following the introduction of the SA-DCP. Only one responder perceived that his workload had slightly increased because the SA-DCP did not support integration with electronic patient record; therefore, patient data had to be transferred manually from one program to another (Textbox 1). However, HCPs reported that in some situations, the SA-DCP brought more flexibility to their work practices regarding patient communication (Textbox 1). For example, it enabled them to respond to patients' DCP messages during nonurgent work times, not only prereserved times. HCPs also reported that the initial guiding session of CPAP therapy went more smoothly for SA-DCP users who had familiarized themselves with the information about CPAP therapy through the SA-DCP (Textbox 1). The interviewed HCPs hoped that patients would make more use of the SA-DCP and its possibilities so its benefits would be better used.

XSL•FO
RenderX

Textbox 1. Themes and perceived barriers and facilitators regarding implementation of the sleep apnea digital care pathway (SA-DCP) according to health care professionals (HCPs).

Use rate of SA-DCP

- Barriers
  - SA-DCP's use rate had been lower than HCPs assumed it would be.
  - Some patients still thought that the only proper contact was personal contact with HCPs.
  - Patients' previous experiences with the need to log in to several digital health care services reduced their motivation to use them.
- Facilitators
  - Reminder text messages about logging into the SA-DCP have been sent to patients since June 2020.

Initial guiding session of continuous positive airway pressure (CPAP) therapy

- Barriers
  - Some SA-DCP users and nonusers still had not familiarized themselves with the prior information about CPAP therapy in advance.
- Facilitators
  - The guidance went more smoothly for SA-DCP users who familiarized themselves with the prior information about CPAP therapy.
  - From the patients' perspective, the instructional videos available in the SA-DCP were perceived as useful and clear.

Patients' communication practices with HCP

- Barriers
  - There were still a lot of phone calls.
  - HCPs also had to be reminded that they should not always call patients in connection with treatment controls but send a message via the SA-DCP.
- Facilitators
  - The SA-DCP gives patients more flexibility to contact HCPs regardless of time and place. For example, the patient may be in a location where they cannot answer the HCP's phone call.
  - HCPs may instruct the patient during a phone call to watch SA-DCP's educational video to get a better understanding of the matter.

Patients' adherence to CPAP therapy

- Barriers
  - There was no clear indication that patients' adherence to CPAP therapy was higher with the introduction of the SA-DCP.
- Facilitators
  - Reports obtained from CPAP devices had increased some patients' adherence to CPAP therapy.

Integration of SA-DCP into existing information and communications technology systems

- Barriers
  - The SA-DCP had to be used in a different web browser than electronic patient record (EPR).
  - Remote monitoring of CPAP devices requires a separate program, and remote monitoring data cannot be viewed via the SA-DCP.
  - Attaching patients to the SA-DCP is laborious and must be done manually by copying patient information from the EPR.
  - Data had to be copied manually from SA-DCP's messages into patients' care plans.

Workload and work practices of HCPs

- Barriers
  - HCPs' workloads did not change with the introduction of the SA-DCP.
- Facilitators

XSL•FO
RenderX

- The SA-DCP brought more flexibility to HCPs' work practices regarding communication with patients.
- The SA-DCP was a good way to deliver the necessary contact information to patients and thereby instruct them to reserve time for the necessary procedures by themselves.

The professionals also brought up ideas for the development of SA-DCP. They hoped that SA-DCP's integration with other information and communications technology (ICT) systems would be improved. One factor that caused a large workload for professionals was arranging appointment times for patients. The time reserved for the patient may not always suit him or her, necessitating a discussion about a more suitable time. If the patient could book appointments through the SA-DCP, it would greatly reduce the professionals' working hours. Two respondents mentioned that in the future, an initial guiding session of CPAP therapy could also be carried out remotely, but this would require that patients for whom this would be suitable should be identified in advance. The professionals hoped that all surveys and measurements made by the patients related to their treatment would be available in an electronic format. The hope was also that the SA-DCP's calendar would automatically remind patients, for example, to renew equipment, giving them more responsibility for managing their own affairs. One respondent wished that instructional videos could be directly linked to SA-DCP's messages so that patients would not have to search for them.

The HCP interviewees perceived digital services in health care as a positive thing. According to them, the services should be easy to use, and the real end users of the services should be included in their development. One respondent believed that patients will use digital health care services more frequently in the future, but such systems are always initially met with resistance. The respondent mentioned that at first, patients in Finland were against e-prescribing and the Patient Data Repository of Kanta Services, but today, such services are commonplace, and people use them smoothly.

Comparison of Characteristics Between SA-DCP Users and SA-DCP Nonusers

According to the patients' survey, there were no statistically significant differences in age, sex, and smoking status between SA-DCP users and nonusers (Table 2). According to the remote monitoring data of CPAP devices, SA-DCP nonusers were older than SA-DCP users (mean 59.1, SD 13.8 vs mean 55.3, SD 10.8; $P<.049$; Table 3). Compared with nonusers, SA-DCP users perceived their own abilities to use ICT to be better (mean 4.2, SD 0.8 vs mean 3.2, SD 1.2; $P<.001$); they used computers, tablets, or smartphones more often (58/58, 100% vs 27/33 81%; overall $P=.002$); and they were more accustomed to using electronic services (mean 4.8, SD 0.5 vs mean 4.1, SD 1.2; $P=.006$; Table 2). There was no statistically significant difference in how regularly SA-DCP users and nonusers used the electronic services (Table 2). SA-DCP users thought that communication about SA-DCP and how to log in had been clear, although SA-DCP nonusers thought that it had not (yes 52/58, 91% vs yes 7/33, 24%; overall $P<.001$; Table 2). Compared with SA-DCP users, SA-DCP nonusers preferred phone calls or physical appointments with HCPs to manage their health-related issues (Table 2). Neither SA-DCP users nor SA-DCP nonusers had any major concerns about the data security and protection of digital health care services (Table 2).

XSL•FO
RenderX

Table 2. Patient responses to the survey.

Characteristics	Use of SA-DCP[a]		*P* value
	Nonusers (n=33)	Users (n=58)	
Age (years), mean (SD)	61.9 (11.6)	57.3 (12.0)	.08
Sex, n (%)			.07
Male	26 (79)	34 (59)	
Female	7 (21)	24 (41)	
Physical training frequency, n (%)			.03
Daily	16 (49)	15 (26)	
Weekly	13 (39)	37 (64)	
Monthly	1 (3)	5 (9)	
Less than monthly	1 (3)	0 (0)	
No physical training	2 (3)	1 (2)	
Smoking, n (%)			>.99
Yes	2 (6)	4 (7)	
No	31 (94)	53 (93)	
Adherence to CPAP[b] therapy (own assessment; Likert scale 1-5), mean (SD)	4.8 (0.5)	4.7 (0.8)	.87
Patients familiar with CPAP therapy before the initial guiding session	16 (49)	43 (74)	.02
The patient became familiar through SA-DCP, n (%)	N/A[c]	29 (67)	__[d]
Average use of the CPAP device per night (hours), mean (SD)	6.3 (1.0)	6.3 (1.3)	.97
Has CPAP therapy helped the patient's sleep apnea?, n (%)			>.99
Yes	28 (85)	49 (85)	
No	0 (0)	0 (0)	
Cannot say	5 (15)	9 (15)	
Information and communication technology skills (own assessment; Likert scale 1-5), mean (SD)	3.2 (1.2)	4.2 (0.8)	<.001
Patient's computer, tablet, or smartphone use, n (%)			.002
Regularly (weekly)	27 (82)	58 (100)	
Randomly (less often than weekly)	4 (12)	0 (0)	
None	2 (6)	0 (0)	
How accustomed is the patient to using electronic services (eg, banking services, appointment services, etc; own assessment)? (Likert scale 1-5), mean (SD)	4.1 (1.2)	4.8 (0.5)	.006
If the patient uses electronic services, how regularly?, n (%)			.55
Daily	20 (69)	44 (76)	
Weekly	7 (24)	11 (19)	
Monthly	1 (3)	3 (5)	
Less often than monthly	1 (3)	0 (0)	
Would the patient choose an electronic service or a phone call as a contact method regarding her or his treatment?, n (%)			<.001
Electronic service	13 (39)	48 (83)	
Phone call	20 (61)	10 (17)	
If the patient could choose either an electronic service (eg, remote consultation) or a physical appointment regarding her or his treatment, which method would she or he prefer?, n (%)			.048
Electronic service	10 (30)	31 (53)	
Physical appointment	23 (70)	27 (46)	

XSL•FO
RenderX

Characteristics	Use of SA-DCP[a]		*P* value
	Nonusers (n=33)	Users (n=58)	
Patient concerns about the data security and protection of digital health care services (Likert scale 1-5), mean (SD)	2.6 (1.2)	2.2 (1.1)	.23
Has communication about SA-DCP and how to log in to it been sufficiently clear?, n (%)			<.001
Yes	7 (24)	52 (91)	
No	22 (76)	5 (9)	
Did SA-DCP increase the patient's adherence to CPAP therapy?, n (%)			
Yes	N/A	42 (72)	
No	N/A	16 (28)	
Did the patient contact HCP[e] during her or his treatment period?, n (%)	17 (52)	33 (57)	.67
Through SA-DCP	N/A	17 (52)	
Through another contact method	17 (100)	16 (48)	
The contact was related to (total), n	26	37	—
Treatment of sleep apnea, n (%)	8 (31)	3 (8)	—
CPAP therapy, n (%)	15 (58)	23 (62)	—
Other issues, n (%)	3 (12)	11 (30)	—
Did patients who contacted HCP get the help they needed?, n (%)			
Yes	16 (94)	31 (94)	.41
Through SA-DCP messaging	N/A	16 (52)	—
Through another contact method	16 (100)	15 (48)	—
Did the patient need to find additional information about his or her treatment without contacting HCPs during the treatment period?, n (%)	9 (27)	23 (40)	.48
The patient got the information she or he needed	8 (89)	21 (91)	>.99
Through SA-DCP	N/A	8 (38)	—
Through another source (internet, patient organizations, etc)	8 (100)	13 (62)	—

[a]SA-DCP: sleep apnea digital care pathway.

[b]CPAP: continuous positive airway pressure.

[c]N/A: not applicable.

[d]Not available.

[e]HCP: health care professional.

XSL•FO
RenderX

Table 3. Remote monitoring data of continuous positive airway pressure (CPAP) devices.

Characteristics	Use of SA-DCP[a]		*P* value
	Nonusers (n=90)	Users (n=80)	
Age (years), mean (SD)	59.1 (13.8)	55.3 (10.8)	.049
Sex, n (%)			.35
Male	48 (53.3)	49 (61.3)	
Female	42 (46.7)	31 (38.8)	
AHI[b] at diagnosis, mean (SD)	32.5 (18.0)	30.6 (18.9)	.51
AHI residual in treatment, mean (SD)	2.4 (2.7)	2.1 (3.9)	.51
Percentage of nights CPAP was used, mean (SD)	92.6 (12.7)	91.5 (18.8)	.66
Hours of CPAP use per night, mean (SD)	6.2 (1.5)	6.1 (1.8)	.69
CPAP device mask leak, mean (SD)	3.1 (6.1)	2.8 (3.2)	.74
CPAP device median pressure, mean (SD)	8.4 (2.2)	7.8 (2.0)	.06

[a]SA-DCP: sleep apnea digital care pathway.

[b]AHI: apnea-hypopnea index.

Patients' Rationales for Using or Not Using SA-DCP

Patients were asked about their rationale for using or not using the SA-DCP (Table 4). SA-DCP users mostly adopted the SA-DCP because they thought that signing up for the SA-DCP was part of their treatment process (42/58, 72%). SA-DCP nonusers did not adopt the SA-DCP mainly because they were unaware of it (15/33, 46%).

XSL•FO
RenderX

Table 4. Patients' rationales for using or not using the sleep apnea digital care pathway (SA-DCP).

Patients' rationales for using or not using the sleep apnea digital care pathway (SA-DCP)	Values, n (%)
Patients' rationales for using the SA-DCP (n=58)	
I thought signing up for SA-DCP was part of my treatment process	42 (72)
It was recommended to me	39 (67)
It allows me to take care of my affairs regardless of time and place	31 (53)
I am very accustomed to using electronic services	25 (42)
I can more easily get information about sleep apnea and its treatment	18 (31)
I prefer to use electronic services for my treatment	15 (26)
I can more easily get information about CPAP[a] therapy	15 (26)
I can take care of things related to my care more safely during the current COVID-19 period	12 (21)
By using SA-DCP, I am more committed to my treatment	8 (14)
Other reasons	2 (3)
Patients' rationales for not using the SA-DCP (n=33)	
I am not aware of SA-DCP	15 (46)
I prefer physical appointments	13 (39)
I prefer phone calls	10 (30)
I am aware of SA-DCP, but I forgot to log in	9 (27)
I do not know how to use electronic services	7 (21)
I do not want to use electronic services	6 (18)
The use of electronic services is generally difficult	6 (18)
I do not receive personal help through electronic services	5 (15)
Other reasons	4 (12)
I am concerned about the data security and protection of electronic services	3 (9)
My sleep apnea treatment and CPAP therapy are balanced, so I do not need to contact health care professionals through any communication channel.	3 (9)
My sleep apnea treatment and CPAP therapy are balanced, so I do not need additional information through any communication channel.	1 (3)
I do not feel the need to log into SA-DCP as part of my CPAP therapy	1 (3)

[a]CPAP: continuous positive airway pressure.

Patients' Prefamiliarization With CPAP Therapy Before the Initial Guiding Session

A larger proportion of SA-DCP users had familiarized themselves with prior information about CPAP therapy before the initial guiding session of CPAP therapy than SA-DCP nonusers (43/58, 74% vs 16/33, 49%; *P*=.02; Table 2). Among the 48 SA-DCP users who familiarized themselves with information about CPAP therapy beforehand, 29 (67%) performed it through the SA-DCP (Table 2). Most SA-DCP nonusers (6/16, 38%) said they had received the preliminary information from a spouse or a relative who had already used a CPAP device. The other sources of information for both groups were the internet (6/59, 10%), private health care providers (3/59, 5%), primary health care units (2/59, 3%), and the Duodecim medical information database (2/59, 3%). SA-DCP users also received information from occupational health care units (2/43, 5%) and the Facebook sleep apnea support group (2/43, 5%). Correspondingly, SA-DCP nonusers received information from specialized care units (2/16, 13%), research articles (1/16, 6%), and AEC (1/16, 6%).

The initial guidance sessions of CPAP therapy were carried out with small groups of patients (4-8 patients at a time). According to HCPs, the initial guiding sessions were smoother for patients who had already familiarized themselves with prior information about CPAP therapy through the SA-DCP (Textbox 1). The problem was that many patients still did not have prior information about CPAP therapy; therefore, most of the initial guiding sessions had to be implemented according to their needs. According to the HCPs, patients found the instructional videos available in the SA-DCP to be useful and clear (Textbox 1). Patients were also instructed to familiarize themselves with them and other information material found on the SA-DCP even after the sessions if they had further questions.

XSL•FO
RenderX

Patients' Information Needs About Sleep Apnea and CPAP Therapy

SA-DCP includes electronic messaging functionality between patients and HCPs and information about sleep apnea, its treatment, and CPAP therapy. According to the survey responses of SA-DCP users, most patients looked for information about the SA-DCP, sleep apnea, self-treatment of sleep apnea, and cleaning and maintenance of the CPAP device (Table 5). The messaging functionality of the SA-DCP and its "frequently asked questions" function were not widely used; only 38% (22/58) of SA-DCP users used them (Table 5).

Table 5. Functionalities of the sleep apnea digital care pathway (SA-DCP) used by patients according to the SA-DCP users survey (N=58).

	Values, n (%)
The number of patients who familiarized themselves with the following information materials	
Welcome to SA-DCP	55 (95)
Sleep apnea	45 (78)
Cleaning and maintenance of the CPAP[a] device	43 (74)
Self-treatment of sleep apnea	42 (72)
CPAP therapy	39 (67)
Preparing for the CPAP therapy initial guiding session	38 (67)
Sleep apnea and driving ability	32 (55)
CPAP therapy in unusual everyday situations	28 (48)
Controls, rehabilitation, and social security	25 (43)
Frequently asked questions	22 (38)
The patient has communicated with the HCP[b] in matters related to his or her treatment through the messaging functionality of SA-DCP	22 (38)

[a]CPAP: continuous positive airway pressure.

[b]HCP: health care professional.

During the treatment period, both SA-DCP users and SA-DCP nonusers sought additional information regarding their treatment without contacting HCPs (23/58, 39% vs 9/33, 27%; *P*=.48; Table 2). Among 23 SA-DCP users who sought more information, 8 (35%) performed it through the SA-DCP. The other reported information sources for SA-DCP users were the internet (9/23, 39%), Facebook sleep apnea support group (1/23, 4%), patient organizations (1/23, 4%), rehabilitation (1/23, 4%), and professional education (1/23, 4%). SA-DCP nonusers received additional information from the following sources: the internet (5/9, 56%), Facebook sleep apnea support group (1/9, 11%), scientific articles (1/9, 11%), and information material provided by private health care services providers (1/9, 11%). Most SA-DCP nonusers and SA-DCP users who sought more information about CPAP therapy and sleep apnea received the information they needed (8/9, 89% vs 21/23, 91%; *P*>.99; Table 2).

Patients' Communication Practices With HCPs

During the treatment period, 52% (17/33) of SA-DCP nonusers and 57% (33/58) of SA-DCP users contacted HCPs (Table 2). Among 33 SA-DCP users who contacted HCPs, 17 (52%) used the SA-DCP, and the rest used other contact methods (Table 2). SA-DCP users who preferred contact methods other than SA-DCP messages were older (mean 63.3, SD 9.4 vs mean 55.7, SD 10.8; *P*<.04). Phone calls were the most important form of contact for SA-DCP users (10/33, 30%). SA-DCP nonusers (9/17, 53%) mostly contacted HCP via phone calls. The next most common contact method for both groups was a physical visit to an AEC or a primary health care unit. The contact mostly concerned CPAP therapy; this was the case for 88% (15/17) of SA-DCP nonusers and 70% (23/33) of SA-DCP users (Table 2). Most SA-DCP nonusers and SA-DCP users who contacted HCPs received the help they needed (16/17, 94% vs 31/33, 94%; *P*=.41; Table 2).

According to the phone call register data, the annual number of phone calls to an AEC was still high even after the introduction of SA-DCP (Table 1). An exact comparison of phone calls to AECs per patient between different years could not be made because the number of new CPAP therapies in the OUH varied between different years, and the number of annual phone calls also showed contacts with HCPs from patients whose CPAP therapies had started in previous years (Table 1). The results of 2019 mainly represent a situation in which the SA-DCP was not yet in use at OUH because it was introduced at the very end of 2019. The results of 2021 represent a situation in which the SA-DCP had been in use at OUH for approximately 2 years. HCPs also indicated the same; there was no significant decrease in the number of phone calls, and there were still many phone calls related to CPAP therapy (Textbox 1). HCPs emphasized that they try to guide patients during phone calls to use the SA-DCP more in matters related to their care. Although the patients' affairs were handled mostly with phone calls, the HCPs thought the instructional videos and informational materials included in the SA-DCP were valuable. It was possible to better explain things to patients with them (Textbox 1). For example, HCPs may instruct the patient during a phone call to watch SA-DCP's educational video to get a better understanding of the matter. The HCPs also emphasized that the SA-DCP is a

XSL•FO
RenderX

good way to deliver the necessary contact information to patients, allowing them to reserve time for the necessary procedures themselves (Textbox 1).

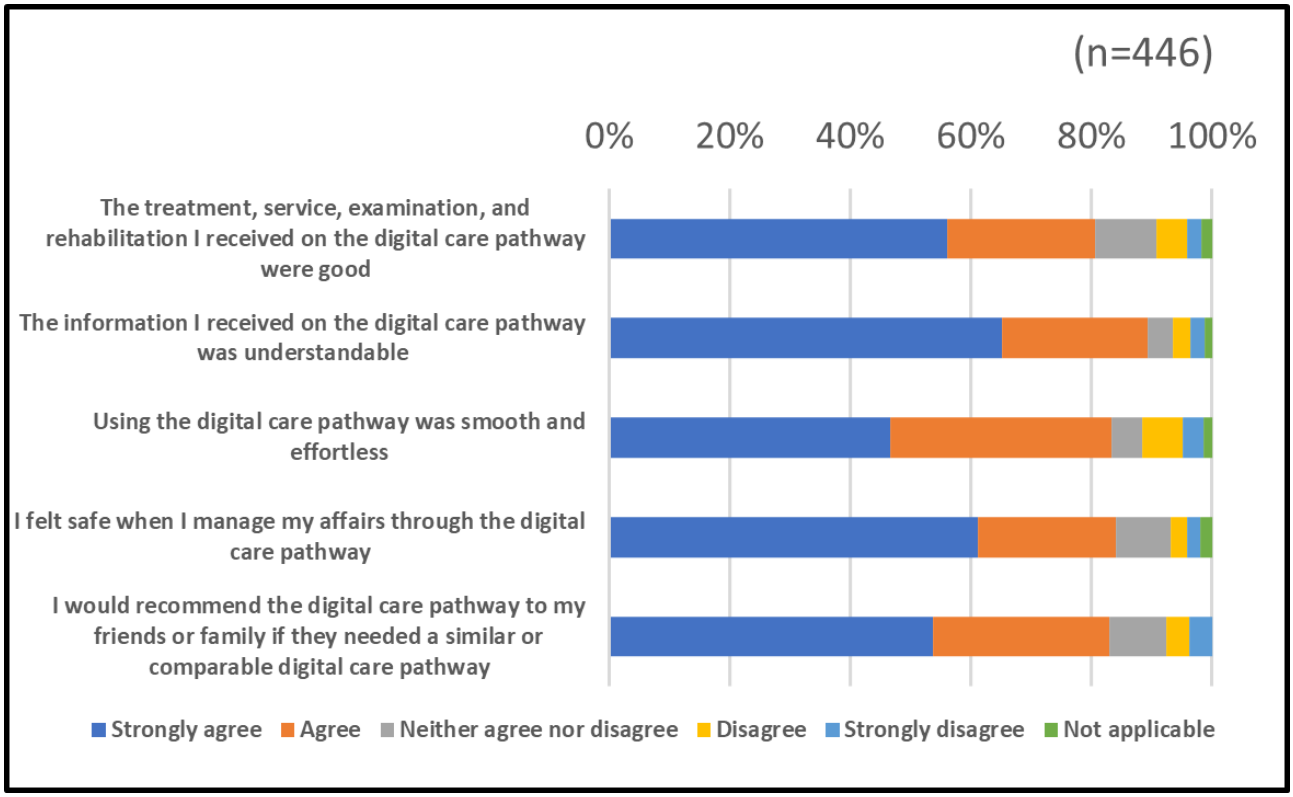
Patient Adherence to CPAP Therapy

According to the patients' responses to the survey and remote monitoring data of CPAP devices, adherence to CPAP therapy was high in both groups (Tables 2 and 3). Both groups used the CPAP device on average for >6 hours per night and on >90% of nights (Tables 2 and 3). On the basis of the patients' own assessments, adherence to CPAP therapy was high in both SA-DCP nonusers (mean 4.8, SD 0.5) and SA-DCP users (mean 4.7, SD 0.8; Table 2). In addition, according to the patients' survey, 72% (42/58) of SA-DCP users reported that SA-DCP had made them more motivated to perform their own CPAP therapy (Table 2). Most patients in both groups believed that CPAP therapy helped them treat sleep apnea (Table 2). The remote monitoring data of CPAP devices showed that CPAP therapy had significantly reduced the number of AHIs for both groups (Table 3).

Patient Feedback About SA-DCP

A total of 446 patients responded to the customer feedback survey; their feedback is shown in Figure 1. Patient feedback on the SA-DCP was generally positive; most of them agreed or strongly agreed with the survey claims (Figure 1). When examining the results, it should be noted that the questions of the patient feedback survey are common to every DCP in the OUH. As the functionalities offered by DCPs vary according to the care chains of different diseases, not all the questions are necessarily valid for every DCP. For example, examinations are not offered through the SA-DCP.

Figure 1. Patient feedback about the sleep apnea digital care pathway.



Of 446 patients, 102 (22.9%) who responded to the survey provided free-form feedback on the SA-DCP. Moreover, 22 patients gave generally positive feedback about the SA-DCP. For the most part, they did not elaborate on their feedback. According to 2 respondents, the possibility to use the services remotely was a good thing, and according to 2 respondents, the SA-DCP was a good and modern service. However, 19 patients thought that they did not need to use the SA-DCP, or that it did not add value to their treatment. Moreover, 11 respondents mentioned that communication and information about the SA-DCP should be improved. According to 9 respondents, the SA-DCP contained good and comprehensive information about sleep apnea and its treatment, as well as CPAP therapy. However, 3 respondents mentioned that although the SA-DCP contained good information, the same information can be found on the internet. With regard to SA-DCP's messaging feature, 5 respondents thought it was a functional solution. Conversely, 9 respondents said that they encountered problems or delays related to messaging and 9 respondents desired new features for the SA-DCP, such as better search functionality. As the information content of SA-DCP was only available in Finnish during the research, some respondents presented English language support as a need for future development. According to 5 responses, SA-DCP's user interface was clear, and its usability was good. In contrast, 4 respondents stated that the user interface could still be improved. Three respondents had technical problems and challenges when using the SA-DCP. Two users reported that the SA-DCP worked well technically. Three respondents said that they would not like to manage their affairs through digital services. Four respondents reported that they had experienced challenges using the SA-DCP, especially in relation to finding their own care path.

XSL•FO
RenderX

Discussion

Principal Findings

This study investigated whether the 3 main goals for introducing the SA-DCP at OUH were achieved. The first aim of introducing the SA-DCP was to shorten the initial guiding sessions of CPAP therapy on the assumption that the patients would have familiarized themselves with prior information about CPAP therapy in advance through the SA-DCP. The second main aim was to reduce the number of patients' phone calls and contacts to HCPs, especially when the information can be found in the SA-DCP. The primary goal of implementing SA-DCP at OUH was to improve patients' adherence to CPAP therapy. However, according to the results of this study, not all the objectives of introducing the SA-DCP were achieved.

On the basis of the HCP's responses to this study, shortening the initial guiding sessions of CPAP therapy had not been fully achieved, although a significantly larger number of SA-DCP users had familiarized themselves with prior information about CPAP therapy compared with SA-DCP nonusers. In this regard, it can be said that SA-DCP has contributed to the better preparation of patients for sessions. The initial guiding sessions were smoother for patients who had already familiarized themselves with prior information regarding CPAP therapy through the SA-DCP. However, many patients still did not have prior information about CPAP therapy; therefore, most sessions had to be implemented according to their needs. Because digital services may require care process changes to get the most out of them, 2 HCPs mentioned that the initial guidance sessions could also be carried out remotely in the future; however, this would require that the patients for whom this procedure would be suitable should be identified in advance [11,12].

Despite previous studies showing that DCPs would make it possible to reduce the number of patient phone calls to HCPs, this did not happen in the case of SA-DCP [37,38]. The annual number of phone calls to an AEC was still high even after the introduction of SA-DCP, according to the phone call register data. As the number of patients' phone calls related to CPAP therapy was still high, HCPs mentioned that it was difficult to assess the actual change in the number of phone calls. However, they perceived that the number of patient calls did not decrease significantly. Previous studies have shown that patients' ability to use electronic services also promotes the use of digital health care services [39,40]. However, Jenssen et al [41] found that despite the regular use of new digital technologies and services such as electronic banking, few of their study participants supported using these tools for communicating with their HCPs. The same behavior pattern can also be observed in the case of SA-DCP. Although SA-DCP users in this study perceived their ability to use ICT to be good and used computers, tablets, or smartphones regularly and were accustomed to using electronic services, only approximately half of them contacted the HCP with SA-DCP messages when needed. Among SA-DCP users, phone calls were the most important other contact method. The notable finding was that SA-DCP users who preferred another contact method were older.

Patient concerns about data security and protection have weakened their willingness to use electronic communication methods in health care [15,42]. On the basis of this study, this would not be an explanatory factor for the low use of SA-DCP messages, as both SA-DCP users and SA-DCP nonusers were not significantly concerned about the data security and protection of digital health care services. Zanaboni and Fagerlund [43] discovered that communicating via electronic tools was less time-consuming from the patient's perspective than communicating via phone calls. However, some participants indicated that the time elapsed to receive a response from the HCP was more important than the time spent using the service itself. Long response times have been seen as one of the most important reasons for patients' dissatisfaction with electronic communication in health care [39,44]. In a Norwegian study, older patients hoped that their electronic messages would be answered the next day at the latest; otherwise, they experienced dissatisfaction with the service [39]. From the patients' point of view, they may perceive that a phone call is a quick and convenient way to handle their health-related matters [45-47]. The fundamental difference is that a phone call involves real-time interactive communication, whereas SA-DCP messages can be defined as asynchronous communication [48]. The patient may ask follow-up questions during the phone call and the HCP can answer them immediately. When using electronic communication tools, there may be delays in answers to questions and possible follow-up questions because of asynchronous communication, as the patient and the HCP may not be dealing with the issue simultaneously [48].

One of the main goals of introducing the SA-DCP was to improve patients' adherence to CPAP therapy. This study showed no statistical difference between SA-DCP users' and nonusers' adherence to CPAP therapy. Adherence to CPAP therapy was high in both groups according to the patients' own estimates and remote monitoring data of CPAP devices. Both groups performed CPAP therapy regularly and reported that it helped them to treat their sleep apnea. In addition, 72% (42/58) of SA-DCP users reported that SA-DCP motivated them to perform their own CPAP therapy. Unfortunately, this study did not ask why the participants felt this. The role of the SA-DCP was to complement CPAP therapy by providing information and an electronic communication channel. It did not include clear mechanisms for influencing patients' behavior related to their own health as digital health interventions typically do, for example, in relation to weight management [49-51]. The CPAP therapy clearly helped the participants in this study to reduce the number of AHIs. Presumably, the biggest motivation for performing CPAP therapy came from alleviating sleep apnea symptoms and not so much from using the SA-DCP; therefore, the SA-DCP was not a significant factor in explaining adherence to CPAP therapy.

This study investigated HCPs' perspectives on the SA-DCP and its usefulness for their work. Although previous studies determined that DCPs could potentially free health care services capacity for other purposes and reduce the workload of HCPs, the results of this study do not support these results in the case of SA-DCP [27,28]. The HCPs who participated in the study were unable to define significant changes in their workload and

XSL•FO
RenderX

work practices after the introduction of SA-DCP. The primary aim of HCPs was for patients to use the SA-DCP more so that its benefits could be better used. Previous studies have highlighted that DCPs can promote work flexibility, for example, by enabling HCPs to respond to patients' DCP messages at nonurgent, not only prereserved times [27,31]. The responses of HCPs in this study pointed out the same. With the help of the DCP, patients can access the information it contains before and after contact with HCPs, thus reducing patient follow-up questions [12]. From this perspective, HCPs felt that educational videos and information materials on SA-DCP were beneficial because, through them, the patients could better understand things. From a technical point of view, the SA-DCP's weak integration with existing ICT systems was seen as one of its key shortcomings and an area for future development. The lack of interoperability with existing ICT systems has been found to weaken the willingness of HCPs to use digital health care services and increase their workload [10,52]. According to the interviewed HCPs, lack of integration reduced the fluency of their work, increased the workload of one responder, and can cause risks from the perspective of information protection and patient safety when patient information is copied manually between different programs.

Digital health care services are intended to help patients become more active actors, more adherent to their own care, and change their behavior in a more favorable direction for their health [49-51,53,54]. Promising results have already been achieved, for example, in treating obesity with the help of digital services [30,51]. In the case of SA-DCP, it was hoped that patients would be active and familiarize themselves with the information contained in it about sleep apnea and CPAP therapy. According to the patient survey, most SA-DCP users have done so. Although most SA-DCP users familiarized themselves with the information in SA-DCP, there was no statistically significant difference in the proportion of SA-DCP users and nonusers who sought additional information about their illness or CPAP therapy. From this perspective, it cannot be said that SA-DCP users are more active actors. It has been established that digital health care services can lower the threshold for patients to contact HCPs [37,54,55]. According to this study, there was no statistically significant difference between the percentage of SA-DCP users and SA-DCP nonusers who contacted HCPs during their treatment period. However, this study did not ask how often the patients contacted the HCPs. On the basis of the results of this study, it seems that patients sought additional information about their illness or contacted HCPs when they had a real need, regardless of the information source or communication method.

Most SA-DCP users thought that the treatment they received through the SA-DCP was good; it was fine technically, a safe service, and the information it contained was clear and understandable. However, some patients still did not use SA-DCP, although the relative number of active SA-DCP users increased during the study period. Lack of digital literacy is one of the barriers to promoting the use of digital health care services. Older adults, in particular, tend to have lower digital literacy than the general population [39,56]. Mannheim et al [40] emphasized in their study that older adults are not a homogeneous group in terms of digital literacy and should also be better included when designing digital health care services [40]. On the basis of the patient survey, there was no statistically significant difference in the age of SA-DCP users and SA-DCP nonusers, but based on remote monitoring data from CPAP devices, SA-DCP nonusers were older. According to this study, SA-DCP nonusers perceived their abilities to use ICT to be worse; they used computers, tablets, or smartphones more rarely and were less accustomed to using electronic services than SA-DCP users. SA-DCP nonusers preferred phone calls or physical appointments to manage their health-related issues with HCPs. The results showed a statistically significant difference in how clearly the patients perceived the communication about SA-DCP. Only 24% (7/38) of SA-DCP nonusers considered communication to be clear, and ignorance of the SA-DCP was the most common reason for them not to use the SA-DCP. After the diagnosis of sleep apnea, the patients received an information letter containing information about the disease and its treatment. This letter also included information on the SA-DCP and how to use it. Did SA-DCP nonusers think the SA-DCP was not adequately explained because they did not want to use digital health care services in the first place and preferred to conduct their health-related issues through phone calls or physical visits? They may not have paid attention to the SA-DCP information letter if they do not typically use or are not willing to use digital health care services or if they perceive they have weak skills in using them.

One of the key findings of this study is that the nonuse of SA-DCP and its functionalities among patients with sleep apnea means that its full potential is not being used. This can be seen, for example, in the initial guiding sessions of CPAP therapy, when some patients still come without prior knowledge. Although the number of SA-DCP users increased during the years covered by this study, not all SA-DCP functionalities were significantly used. In particular, this was reflected in the fact that SA-DCP messages were not widely used; therefore, the number of calls to AECs was not reduced. This study found that lower digital literacy and older age were significant factors in explaining the nonuse of the SA-DCP. Older SA-DCP users more often favored other contact methods, such as phone calls, when contacting HCPs during their treatment period. In the future, special attention should be paid to how digital health care services are designed according to the needs of older adults with weak digital literacy. Care processes should be better adapted to the requirements of digital health care services. Clearly, only the traditional information letter about SA-DCP is not sufficient to encourage all patients to adopt it. If there are challenges in deployment, patients could be more actively encouraged to adopt the SA-DCP and offered support. Previous studies have highlighted that the desire of older adults to use digital health care services can be supported by offering guidance and peer support [39,56]. Studies have also emphasized that both professionals and patients should be closely involved in DCP development to obtain the best benefit and that development should be a continuous process [10,29]. With age, various functional limitations, such as diminished eyesight related to diabetes or deteriorated motor skills owing to rheumatism, can increase and thus make it more difficult to use digital services [57,58]. Therefore, special attention should be

XSL•FO
RenderX

paid to the usability and accessibility of digital services. The real end users should be involved in the design process, as the interviewed HCPs highlighted [22,25,58].

Limitations

Our study had some limitations. Patients with sleep apnea give up CPAP therapy for different reasons, which can bias this study's data regarding patients' adherence to CPAP therapy. Most patients who responded to the survey had continued CPAP therapy for ≥1 year, and remote monitoring data on CPAP devices were collected in connection with 1-year control. Unfortunately, when the study was carried out, no information was available on the proportion of SA-DCP users and SA-DCP nonusers who had discontinued CPAP therapy. This would have provided additional information about patients' adherence to CPAP therapy. Previous results have highlighted that high attrition rates hinder achieving the full benefits of digital health care services. During the implementation of the study, the SA-DCP did not enable the automatic collection of log data on the activity of patients using the SA-DCP, but through the automatic log data, it was only possible to determine that the patient had used the SA-DCP. Therefore, this study did not examine patients' adherence to SA-DCP use, but only whether they had used the service.

On the basis of the study's results, approximately half of SA-DCP users still contacted HCPs in a way other than through SA-DCP messages, although they reported having good digital literacy. Most SA-DCP nonusers also preferred phone calls to contact HCPs. However, in this study, SA-DCP users and SA-DCP nonusers were not asked why some preferred phone calls to contact HCPs instead of electronic messaging. Future research is needed to better understand this behavior pattern. This study did not ask patients how many times they contacted HCPs; it only investigated whether the patients contacted HCPs during their treatment period. Information on the number of contacts would have provided valuable information on whether using the SA-DCP can lower the threshold for contacting HCPs.

One of the goals of the SA-DCP was to increase patients' adherence to CPAP therapy, and most SA-DCP users felt this was the case. Although the results of the survey and the remote monitoring data of the CPAP devices showed that there was no statistically significant difference in adherence to CPAP therapy between the groups, it would have been beneficial to ask SA-DCP users why most of them felt that the SA-DCP had increased their adherence to CPAP therapy. However, this was not investigated in this study. The sample size of the interviewed HCPs was small in this study. However, the answers to the HCPs were mostly consistent. Most of them thought there were no significant changes to their workload and work practices; there were still many phone calls from patients. At the time of writing, the SA-DCP did not enable the automatic collection of log data about the number of electronic messages. If this information had been available, it would have enabled a better comparison between the volumes of phone calls and SA-DCP messages.

Conclusions

According to this study, not all the goals set for introducing the SA-DCP have been achieved. The HCPs who participated in the study could not define significant changes in their workload and work practices after the introduction of SA-DCP. The SA-DCP has brought more flexibility to HCPs' work practices regarding patient communication. Despite using SA-DPC, some patients still wanted to communicate with HCPs by phone. Adherence to CPAP therapy was high in both SA-DCP users and nonusers. Patients' lower digital literacy and older age were the most significant factors explaining the nonuse of the SA-DCP. In the future, more attention should be paid to how these user groups should be considered in the design and introduction of the DCPs.

Acknowledgments

This research was funded by the Strategic Research Council (336278 and 358417). The authors are grateful to the study participants who generously gave their time to share their views on the sleep apnea digital care pathway. The authors would also like to thank the health care professionals from Oulu University Hospital who helped in the practical implementation of the study.

Authors' Contributions

All authors participated sufficiently in the work to take public responsibility for the appropriate portions of the content. JH, TH, and JR were responsible for study conception and design. JH, HM, PL, and TH performed data acquisition. JH analyzed and interpreted the data. JH, TH, MT, PL, and JR drafted the manuscript.

Conflicts of Interest

PL and HM have been involved in the national and regional development of the digital care pathways of Health Village. All other authors declare no other conflicts of interest.

References

1. Palomäki M, Saaresranta T, Anttalainen U, Partinen M, Keto J, Linna M. Multimorbidity and overall comorbidity of sleep apnoea: a Finnish nationwide study. ERJ Open Res 2022 Apr;8(2):1-12 [FREE Full text] [doi: 10.1183/23120541.00646-2021] [Medline: 35677395]
2. Saaresranta T, Anttalainen U, Brander P, Helin-Salmivaara A, Hiltunen L, Himanen SL, et al. Uniapnea (obstruktiivinen uniapnea aikuisilla) [Update on current care guidelines: obstructive sleep apnoea syndrome in adults]. Lääketieteellinen

XSL•FO
RenderX

Aikakauskirja Duodecim 2010;126(22):2633-2635 [FREE Full text] [doi: 10.1183/9781849840248.013911] [Medline: 21188878]
3. Myllylä M, Anttalainen U, Saaresranta T, Laitinen T. Motor vehicle accidents in CPAP-compliant obstructive sleep apnea patients-a long-term observational study. Sleep Breath 2020 Sep;24(3):1089-1095 [FREE Full text] [doi: 10.1007/s11325-020-02023-2] [Medline: 32060778]
4. Giles TL, Lasserson TJ, Smith BJ, White J, Wright J, Cates CJ. Continuous positive airways pressure for obstructive sleep apnoea in adults. Cochrane Database Syst Rev 2006 Jul 19(3):CD001106. [doi: 10.1002/14651858.CD001106.pub3] [Medline: 16855960]
5. Mattila T, Hasala H, Kreivi HR, Avellan-Hietanen H, Bachour A, Herse F, et al. Changes in the societal burden caused by sleep apnoea in Finland from 1996 to 2018: a national registry study. Lancet Reg Health Eur 2022 May;16:100338 [FREE Full text] [doi: 10.1016/j.lanepe.2022.100338] [Medline: 35252943]
6. Hanlon P, Daines L, Campbell C, McKinstry B, Weller D, Pinnock H. Telehealth interventions to support self-management of long-term conditions: a systematic metareview of diabetes, heart failure, asthma, chronic obstructive pulmonary disease, and cancer. J Med Internet Res 2017 May 17;19(5):e172 [FREE Full text] [doi: 10.2196/jmir.6688] [Medline: 28526671]
7. Liljamo P, Wahlberg J, Mikkonen H, Reponen J. A digital care pathway to access healthcare without time and place restrictions. Stud Health Technol Inform 2021 Dec 15;284:163-165. [doi: 10.3233/SHTI210690] [Medline: 34920495]
8. eHealth. World Health Organization. URL: http://www.emro.who.int/health-topics/ehealth/ [accessed 2023-03-23]
9. O'Reilly-Jacob M, Mohr P, Ellen M, Petersen C, Sarkisian C, Attipoe S, et al. Digital health and low-value care. Healthc (Amst) 2021 Jun;9(2):100533 [FREE Full text] [doi: 10.1016/j.hjdsi.2021.100533] [Medline: 33714891]
10. Kujala S, Ammenwerth E, Kolanen H, Ervast M. Applying and extending the FITT framework to identify the challenges and opportunities of successful eHealth services for patient self-management: qualitative interview study. J Med Internet Res 2020 Aug 12;22(8):e17696 [FREE Full text] [doi: 10.2196/17696] [Medline: 32784175]
11. Koivuluoma M, Haverinen J, Reponen J. Digital health services as part of care processes. FinJeHeW 2022 Nov 07;14(3):284-298. [doi: 10.23996/fjhw.111788]
12. Heijsters F, Santema J, Mullender M, Bouman MB, Bruijne MD, van Nassau F. Stakeholders barriers and facilitators for the implementation of a personalised digital care pathway: a qualitative study. BMJ Open 2022 Nov 23;12(11):e065778 [FREE Full text] [doi: 10.1136/bmjopen-2022-065778] [Medline: 36418140]
13. Saranto K, Kivekäs E, Kuosmanen P, Kinnunen UM. Electronic health services in the patients' daily activities - willingness to use health village services. Stud Health Technol Inform 2018;247:586-590. [Medline: 29678028]
14. Ciere Y, van der Vaart R, van der Meulen-De Jong AE, Maljaars PW, van Buul AR, Koopmans JG, et al. Implementation of an eHealth self-management care path for chronic somatic conditions. Clin eHealth 2019;2:3-11. [doi: 10.1016/j.ceh.2019.04.001]
15. Deng Z, Hong Z, Ren C, Zhang W, Xiang F. What predicts patients' adoption intention toward mHealth services in China: empirical study. JMIR Mhealth Uhealth 2018 Aug 29;6(8):e172 [FREE Full text] [doi: 10.2196/mhealth.9316] [Medline: 30158101]
16. Jääskelä J, Haverinen J, Kaksonen R, Reponen J, Halunen K, Tokola T. Digi-HTA, assessment framework for digital healthcare services: information security and data protection in health technology – initial experiences. FinJeHeW 2022;14(1):19-30 [FREE Full text] [doi: 10.23996/fjhw.111776]
17. Pouls BP, Vriezekolk JE, Bekker CL, Linn AJ, van Onzenoort HA, Vervloet M, et al. Effect of interactive eHealth interventions on improving medication adherence in adults with long-term medication: systematic review. J Med Internet Res 2021 Jan 08;23(1):e18901 [FREE Full text] [doi: 10.2196/18901] [Medline: 33416501]
18. Aardoom JJ, Loheide-Niesmann L, Ossebaard HC, Riper H. Effectiveness of eHealth interventions in improving treatment adherence for adults with obstructive sleep apnea: meta-analytic review. J Med Internet Res 2020 Feb 18;22(2):e16972 [FREE Full text] [doi: 10.2196/16972] [Medline: 32130137]
19. Donkin L, Christensen H, Naismith SL, Neal B, Hickie IB, Glozier N. A systematic review of the impact of adherence on the effectiveness of e-therapies. J Med Internet Res 2011 Aug 05;13(3):e52 [FREE Full text] [doi: 10.2196/jmir.1772] [Medline: 21821503]
20. Manwaring JL, Bryson SW, Goldschmidt AB, Winzelberg AJ, Luce KH, Cunning D, et al. Do adherence variables predict outcome in an online program for the prevention of eating disorders? J Consult Clin Psychol 2008 Apr;76(2):341-346. [doi: 10.1037/0022-006X.76.2.341] [Medline: 18377129]
21. Li S, Cui G, Yin Y, Xu H. Associations between health literacy, digital skill, and eHealth literacy among older Chinese adults: a cross-sectional study. Digit Health 2023;9:20552076231178431 [FREE Full text] [doi: 10.1177/20552076231178431] [Medline: 37256010]
22. Haverinen J, Keränen N, Falkenbach P, Maijala A, Kolehmainen T, Reponen J. Digi-HTA: health technology assessment framework for digital healthcare services. FinJeHeW 2019 Nov 02;11(4):326-341. [doi: 10.23996/fjhw.82538]
23. Haverinen J, Keränen N, Tuovinen T, Ruotanen R, Reponen J. National development and regional differences in eHealth maturity in Finnish public health care: survey study. JMIR Med Inform 2022 Aug 12;10(8):e35612 [FREE Full text] [doi: 10.2196/35612] [Medline: 35969462]

XSL•FO
RenderX

24. Ruotanen R, Kangas M, Tuovinen T, Keränen N, Haverinen J, Reponen J. Finnish e-health services intended for citizens – national and regional development. FinJeHeW 2021 Oct 25;13(3):283-301 [FREE Full text] [doi: 10.23996/fjhw.109778]
25. Haverinen J, Turpeinen M, Falkenbach P, Reponen J. Implementation of a new Digi-HTA process for digital health technologies in Finland. Int J Technol Assess Health Care 2022 Aug 19;38(1):e68. [doi: 10.1017/S0266462322000502] [Medline: 35983625]
26. Digital care pathways. Health Village. 2022. URL: https://www.terveyskyla.fi/en/mypath/digital-care-pathways [accessed 2023-04-01]
27. Tuomikoski K, Liljamo P, Reponen J, Kanste O. The effects of digital care pathways on the healthcare professionals: processes in specialized health care. FinJeHeW 2022 Nov 07;14(3):326-338 [FREE Full text] [doi: 10.23996/fjhw.112648]
28. Väätäinen S, Soini E, Arvonen S, Suvitie P. Cost-benefit predictions of the Virtual Hospital 2.0 digitalized secondary care path for women with endometriosis in Turku area and in Finland. In: Proceedings of the 24th Finnish National Conference on Telemedicine and eHealth. 2019 Nov Presented at: eHealth '19; April 2-3, 2019; Kuopio, Finland. [doi: 10.1016/j.jval.2019.09.1200]
29. Liljamo P, Säilynoja H, Tuomikoski K, Henner A, Koivunen K. Patient-centered development of a digital care pathway for arrhythmia patients. Stud Health Technol Inform 2020 Nov 23;275:127-131. [doi: 10.3233/SHTI200708] [Medline: 33227754]
30. Kupila SK, Venäläinen MS, Suojanen LU, Rosengård-Bärlund M, Ahola AJ, Elo LL, et al. Weight loss trajectories in healthy weight coaching: cohort study. JMIR Form Res 2022 Mar 09;6(3):e26374 [FREE Full text] [doi: 10.2196/26374] [Medline: 35262494]
31. Niemi A, Hupli MI, Koivunen M. The use of electronic communication for patient - professional interaction – nursing staff's point of view. FinJeHeW 2016 Mar;8(4):200-215. [doi: 10.1111/jan.12529] [Medline: 25250739]
32. OYS Uniapnea digihoitopolku. Health Village. 2022. URL: https://www.terveyskyla.fi/omapolku/digihoitopolut#termid=95 [accessed 2023-04-01]
33. Registration and login instructions. Health Village. 2022. URL: https://www.terveyskyla.fi/en/mypath/registration-and-login-instructions [accessed 2023-08-22]
34. Elo S, Kyngäs H. The qualitative content analysis process. J Adv Nurs 2008 Apr;62(1):107-115. [doi: 10.1111/j.1365-2648.2007.04569.x] [Medline: 18352969]
35. Sarja J. Not all numerical data leads to quantitative research: introduction to qualitative quantification analysis. Int J Quant Qual Res Methods 2020;8(2):19-24 [FREE Full text]
36. Ihmiseen kohdistuvan tutkimuksen eettiset periaatteet ja ihmistieteiden eettinen ennakkoarviointi Suomessa. Tutkimuseettinen Neuvottelukunta. 2019. URL: https://tenk.fi/sites/default/files/2021-01/Ihmistieteiden_eettisen_ennakkoarvioinnin_ohje_2020.pdf [accessed 2023-04-01]
37. Laukka E, Huhtakangas M, Heponiemi T, Kujala S, Kaihlanen A, Gluschkoff K, et al. Health care professionals' experiences of patient-professional communication over patient portals: systematic review of qualitative studies. J Med Internet Res 2020 Dec 08;22(12):e21623 [FREE Full text] [doi: 10.2196/21623] [Medline: 33289674]
38. Miller Jr DP, Latulipe C, Melius KA, Quandt SA, Arcury TA. Primary care providers' views of patient portals: interview study of perceived benefits and consequences. J Med Internet Res 2016 Jan 15;18(1):e8 [FREE Full text] [doi: 10.2196/jmir.4953] [Medline: 26772771]
39. Kristiansen E, Atherton H, Austad B, Bergmo T, Norberg BL, Zanaboni P. Older patients' experiences of access to and use of e-consultations with the general practitioner in Norway: an interview study. Scand J Prim Health Care 2023 Mar;41(1):33-42 [FREE Full text] [doi: 10.1080/02813432.2022.2161307] [Medline: 36592342]
40. Mannheim I, Schwartz E, Xi W, Buttigieg SC, McDonnell-Naughton M, Wouters EJ, et al. Inclusion of older adults in the research and design of digital technology. Int J Environ Res Public Health 2019 Oct 02;16(19):3718 [FREE Full text] [doi: 10.3390/ijerph16193718] [Medline: 31581632]
41. Jenssen BP, Mitra N, Shah A, Wan F, Grande D. Using digital technology to engage and communicate with patients: a survey of patient attitudes. J Gen Intern Med 2016 Jan;31(1):85-92 [FREE Full text] [doi: 10.1007/s11606-015-3517-x] [Medline: 26385117]
42. Ye J, Rust G, Fry-Johnson Y, Strothers H. E-mail in patient-provider communication: a systematic review. Patient Educ Couns 2010 Aug;80(2):266-273 [FREE Full text] [doi: 10.1016/j.pec.2009.09.038] [Medline: 19914022]
43. Zanaboni P, Fagerlund AJ. Patients' use and experiences with e-consultation and other digital health services with their general practitioner in Norway: results from an online survey. BMJ Open 2020 Jun 17;10(6):e034773 [FREE Full text] [doi: 10.1136/bmjopen-2019-034773] [Medline: 32554721]
44. Liederman EM, Morefield CS. Web messaging: a new tool for patient-physician communication. J Am Med Inform Assoc 2003;10(3):260-270 [FREE Full text] [doi: 10.1197/jamia.M1259] [Medline: 12626378]
45. Lin C, Wittevrongel L, Moore L, Beaty BL, Ross SE. An internet-based patient-provider communication system: randomized controlled trial. J Med Internet Res 2005 Aug 05;7(4):e47 [FREE Full text] [doi: 10.2196/jmir.7.4.e47] [Medline: 16236699]
46. Kuusisto H, Keränen T, Oksanen K, Saranto K. Neurological referral patients are satisfied with telemedicine consultation. Suomen lääkärilehti 2013;68(9):667-671.

XSL•FO
RenderX

47. Car J, Sheikh A. Telephone consultations. BMJ 2003 May 03;326(7396):966-969 [FREE Full text] [doi: 10.1136/bmj.326.7396.966] [Medline: 12727771]
48. Mechanic OJ, Persaud Y, Kimball AB. Telehealth systems. In: StatPearls [Internet]. Treasure Island, FL: StatPearls Publishing; 2022.
49. Voorheis P, Bhuiya AR, Kuluski K, Pham Q, Petch J. Making sense of theories, models, and frameworks in digital health behavior change design: qualitative descriptive study. J Med Internet Res 2023 Mar 15;25:e45095 [FREE Full text] [doi: 10.2196/45095] [Medline: 36920442]
50. Murray E, Hekler EB, Andersson G, Collins LM, Doherty A, Hollis C, et al. Evaluating digital health interventions: key questions and approaches. Am J Prev Med 2016 Nov;51(5):843-851 [FREE Full text] [doi: 10.1016/j.amepre.2016.06.008] [Medline: 27745684]
51. Teeriniemi AM, Salonurmi T, Jokelainen T, Vähänikkilä H, Alahäivälä T, Karppinen P, et al. A randomized clinical trial of the effectiveness of a web-based health behaviour change support system and group lifestyle counselling on body weight loss in overweight and obese subjects: 2-year outcomes. J Intern Med 2018 Nov 04;284(5):534-545 [FREE Full text] [doi: 10.1111/joim.12802] [Medline: 29974563]
52. Nazi KM. The personal health record paradox: health care professionals' perspectives and the information ecology of personal health record systems in organizational and clinical settings. J Med Internet Res 2013 Apr 04;15(4):e70 [FREE Full text] [doi: 10.2196/jmir.2443] [Medline: 23557596]
53. Blasiak A, Sapanel Y, Leitman D, Ng WY, De Nicola R, Lee VV, et al. Omnichannel communication to boost patient engagement and behavioral change with digital health interventions. J Med Internet Res 2022 Nov 16;24(11):e41463 [FREE Full text] [doi: 10.2196/41463] [Medline: 36383427]
54. Sasseville M, LeBlanc A, Boucher M, Dugas M, Mbemba G, Tchuente J, et al. Digital health interventions for the management of mental health in people with chronic diseases: a rapid review. BMJ Open 2021 Apr 05;11(4):e044437 [FREE Full text] [doi: 10.1136/bmjopen-2020-044437] [Medline: 33820786]
55. Das A, Faxvaag A, Svanæs D. The impact of an eHealth portal on health care professionals' interaction with patients: qualitative study. J Med Internet Res 2015 Nov 24;17(11):e267 [FREE Full text] [doi: 10.2196/jmir.4950] [Medline: 26601678]
56. Schreurs K, Quan-Haase A, Martin K. Problematizing the digital literacy paradox in the context of older adults' ICT use: aging, media discourse, and self-determination. Can J Commun 2017 May 23;42(2):359-377. [doi: 10.22230/cjc.2017v42n2a3130]
57. Wildenbos GA, Peute L, Jaspers M. Aging barriers influencing mobile health usability for older adults: a literature based framework (MOLD-US). Int J Med Inform 2018 Jun;114:66-75. [doi: 10.1016/j.ijmedinf.2018.03.012] [Medline: 29673606]
58. Gilbert BJ, Goodman E, Chadda A, Hatfield D, Forman DE, Panch T. The role of mobile health in elderly populations. Curr Geri Rep 2015 Aug 21;4(4):347-352. [doi: 10.1007/s13670-015-0145-6]

Abbreviations

AEC: assistive equipment center
AHI: apnea-hypopnea index
CPAP: continuous positive airway pressure
DCP: digital care pathway
HCP: health care professional
ICT: information and communications technology
OUH: Oulu University Hospital
SA-DCP: sleep apnea digital care pathway

Edited by A Kushniruk; submitted 02.04.23; peer-reviewed by S Kupila, J Roland, Y Sapanel; comments to author 23.07.23; revised version received 31.08.23; accepted 20.01.24; published 22.02.24.

Please cite as:
Haverinen J, Harju T, Mikkonen H, Liljamo P, Turpeinen M, Reponen J
Digital Care Pathway for Patients With Sleep Apnea in Specialized Care: Mixed Methods Study
JMIR Hum Factors 2024;11:e47809
URL: https://humanfactors.jmir.org/2024/1/e47809
doi:10.2196/47809
PMID:38386368

XSL•FO
RenderX

©Jari Haverinen, Terttu Harju, Hanna Mikkonen, Pia Liljamo, Miia Turpeinen, Jarmo Reponen. Originally published in JMIR Human Factors (https://humanfactors.jmir.org), 22.02.2024. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on https://humanfactors.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

User Experience Evaluation of a Spinal Surgery Robot: Workload, Usability, and Satisfaction Study

Hyeonkyeong Choi[1], BSPH; Seunghee Kim[1], BSPH; Wonseuk Jang[1,2], PhD

[1]Department of Medical Device Engineering and Management, Yonsei University College of Medicine, Seoul, Republic of Korea

[2]Medical Device Usability Research Center, Gangnam Severance Hospital, Yonsei University College of Medicine, Seoul, Republic of Korea

Corresponding Author:
Wonseuk Jang, PhD
Department of Medical Device Engineering and Management
Yonsei University College of Medicine
20 63rd Road, Unju-ro, Gangnam-gu
Seoul, 06229
Republic of Korea
Phone: 82 1091485404
Fax: 82 02 2019 5442
Email: WS.JANG@yuhs.ac

Abstract

Background: Robotic spine surgery has continued to evolve since its US Food and Drug Administration approval in 2004, with products now including real-time video guidance and navigation during surgery. As the market for robotic surgical devices evolves, it is important to consider usability factors.

Objective: The primary objective of this study was to determine the user experience of a surgical-assistive robotic device. The secondary objective was to evaluate workload, usability, the After-Scenario Questionnaire (ASQ), and the System Usability Scale (SUS). In addition, this study compares the workload, usability, and satisfaction survey of the device among different occupational groups using the device.

Methods: Doctors (n=15) and nurses (n=15), the intended users of the surgical assistant robot, participated in the usability evaluation. Participants performed essential scenarios for the surgical assistant robot and provided scenario-specific satisfaction (ASQ), workload (NASA Task Load Index), and usability (SUS) scores.

Results: Both doctors and nurses had task success rates of 85% or higher for each scenario. ASQ results showed that both doctors and nurses were least satisfied with ease of completing the task of registration (group 1: mean 4.73, SD 1.57 and group 2: mean 4.47, SD 1.8), amount of time it took (group 1: mean 4.47, SD 1.63 and group 2: mean 4.40, SD 2.09), and support information satisfaction (group 1: mean 5.13, SD 1.50 and group 2: mean 5.13, SD 1.89). All participants had low workloads, and the overall Task Load Index score had a *P* value of .77, which is greater than .05. The SUS results showed that the overall usability mean for doctors was 64.17 (SD 16.52) and the mean for nurses was 61.67 (SD 19.18), with a *P* value of .84, which is greater than .05, indicating no difference between the 2 groups.

Conclusions: In this study, doctors and nurses evaluated the interaction of the device in a simulated environment, the operating room. By evaluating the use experience and usability of the device with real intended users, we can develop a more effective and convenient user interface.

(JMIR Hum Factors 2024;11:e54425) doi:10.2196/54425

KEYWORDS

robot spine surgery; usability; satisfaction; System Usability Scale; surgical navigation systems; robotics; surgery; neurosurgery

Introduction

Background

Spine surgery is used to treat degenerative diseases and deformities of the spine, with 45 million surgeries performed annually in the United States [1]. The use of robotic-assisted navigation is increasing as the number of patients undergoing lumbar spinal fixation increases [2]. Spine surgery typically involves 7 people in the operating room, with an operator surgeon, a surgical first assistant (who may be a doctor or

XSL•FO
RenderX

physician assistant nurse, depending on operating room staffing), and scrub nurse in the sterile area and a circulating nurse and radiologist in the nonsterile area. Nonoperative personnel include an anesthesiologist and an anesthesiologist assistant. The use of robotics in spine surgery is usually reserved for difficult anatomical areas where it is difficult to fix screws blindly. Spinal fusion surgery is the insertion and fixation of pedicle screws into the vertebrae to eliminate pain by preventing movement between vertebrae [3]. It is also used for quick insertion in severe scoliosis, collapsed vertebrae, or long-level fusion in patients with difficult anatomy, usually at the iliac screw, C1, C2, C7, T1, and T2, or for other reasons. This is usually used for kyphosis and scoliosis correction.

Robot Spine Surgery

Surgical navigation systems are used to plan the procedure and guide the surgeon in inserting the screws [1]. Robot spine surgery is popularly used to increase the accuracy of inserting screws in the spine, and the first robot used in spine surgery was the Spine Assist (Mazor Surgical Technologies), which received Food and Drug Administration clearance in 2004 [4]. The third-generation Mazor X system was cleared by the Food and Drug Administration in 2016 and, compared to previous generations, has a robotic arm that is attached to the patient's body and can be viewed through a camera to ensure that the screws are inserted and the robotic arm is moving well [5]. The Mazor X Stealth Edition technology, which adds real-time image guidance and navigation during surgery, was cleared in 2019 and combines the best of both worlds: traditional spinal robotic surgery guidance and real-time software confirmation [5,6]. Figure 1 shows the evaluation device, which consists of a robotic arm, main console, and optional staff console, and is manufactured in South Korea. Like the Mazor X Stealth Edition technology, this product is capable of real-time image guidance and navigation during surgery.

Figure 1. The CUVIS spine robot system: robotic arm (left), main console with navigation optical infrared tracking camera (middle), and the control workstation (right). OTS: optical tracking system.

Usability

According to IEC (International Electrotechnical Commission) 62366-1 [7], usability has the following meaning: a "characteristic of the user interface that facilitates use and thereby establishes effectiveness, efficiency and user satisfaction in the intended use environment." ISO (International Organization for Standardization) 9241-210 [8] defines usability as the "extent to which a system, product or service can be used by specified users to achieve specified goals with effectiveness, efficiency and satisfaction in a specified context of use." These documents demonstrate that good usability in medical device design is essential to preventing error-related risks.

Evaluating usability focuses on determining whether the test device is easy for users to use. To evaluate the user experience of the device, we used usability tests and surveys for effectiveness, efficiency, satisfaction, and workload. The usability test was mainly used to identify use errors and efficiency, while the After-Scenario Questionnaire (ASQ) was conducted to evaluate the satisfaction of each task. NASA Task Load Index (NASA TLX) was used to measure the workload of the device, and System Usability Scale (SUS) was used to

XSL•FO
RenderX

check the overall system usability of the device. Both ASQ and SUS identified the satisfaction and efficiency of the device, but in this study, ASQ identified the efficiency and satisfaction of each scenario, while SUS evaluated the satisfaction and efficiency of the device in the overall workflow.

Methods

Recruitment

We recruited 30 medical staff from Severance Hospital in South Korea. There were 15 doctors and 15 nurses. The intended users of the device are doctors and operating room nurses. Due to the different tasks that doctors and nurses have to perform when using the device overall, both groups were selected to participate in the usability test. We recruited through recommendations from colleagues and notice on the bulletin board at the Future Medicine Research Center at Gangnam Severance Hospital. Doctors and nurses who have experience using a robotic surgical device or navigation system were selected. For doctors, we selected those with the necessary knowledge of spine surgery, and for nurses, we selected those with experience in the operating room. However, those who had worked in the operating room for less than 1 year were excluded. After confirming these inclusion and exclusion criteria, the screening was conducted.

Testing Procedure

One participant per session participated in the usability test, with the participant completing the assessment in a sequence guided by a facilitator, and an observer in the observation room videotaping the assessment and completing an observation sheet. A total of 2 moderators and 2 observers participated in the evaluation, with one of the moderators acting as a nurse if a doctor participated in the evaluation, and one of the moderators acting as a doctor if a nurse participated in the evaluation. In addition, observers were used to reduce bias by having 2 observers observe the usability test to ensure that one person's opinion was not biased.

The facilitator introduced the participants to the usability test, obtained their informed consent, and trained them on the device for 20 minutes. Participants were allowed to interact with the device as much time as they needed. Afterward, 15 minutes were allowed between the training and evaluation to ensure that the training did not directly influence the evaluation [9]. Doctors and nurses were given different tasks because of the different job duties they do when operating with assessment devices. The tasks given to nurses focus on doctors' instructions from preoperative preparation to surgery, while doctors focus on the surgery itself rather than preparing devices.

Participants then completed the evaluation for 40 minutes, with 8 scenarios (24 tasks) for doctors and 12 scenarios (41 tasks) for nurses. Doctors were asked to complete the following scenarios: preparation for use, preplanning of surgery, fixation of patient marker, scan, registration, verification, revisions of surgical planning, and navigation; nurses were asked to complete the following scenarios: preparation for use, system operation, initializing manipulator, drape, preparation of surgery, scan, registration, verification, planning of surgery, navigation, use of the emergency stop switch, and cleaning up after surgery. Tasks for each scenario are shown in Tables 1 and 2.

XSL•FO
RenderX

Table 1. Use scenarios of doctors.

Use scenario/task number	Task description
Preparation for use	
Task 1	Check the contents related to the emergency button, foot switch, and foot jamming in the user manual.
Task 2	Check the hand jamming label on the robotic arm manipulator.
Preplanning of surgery	
Task 3	Select the spine level as follows: • L4-L, L4-R • L5-L, L5-R
Task 4	After loading the first CT[a] data, check the CT data.
Task 5	Create an implant screw insertion path for target L4-L and L4-R and change the insertion path by moving the screw in the MPR[b] view.
Task 6	Create an implant screw insertion path for target L5-L and L5-R and change the insertion path using the arrow.
Task 7	Check for collision between each screw.
Fixation of patient marker	
Task 8	Attach the patient marker to the patient.
Scan	
Task 9	Attach the registration tool adapter to the end effector.
Task 10	After activating the hand guide function by pressing the AP[c] button, change the position of the end effector according to the guidance on the pop-up window.
Task 11	For C-Arm scan, attach the source calibrator to the end effector in the direction of AP and move the end effector to enable tracking by OTS[d].
Task 12	Check if the ROI[e] includes the calibration marker, and the position and direction of the letters “R,” “G,” and “J” match the image, and then check pass or fail of registration.
Registration	
Task 13	Perform segmentation to distinguish the surgical target in the image. • L4-L, L4-R • L5-L, L5-R
Task 14	Perform labeling to assign target level information of ROI of 3D image and 2D image segmented for each spine level.
Task 15	Adjust the ROI box so that the ROI of target L4 covers all the L4 vertebra area.
Task 16	Perform 2D and 3D image registration for each spine level.
Verification	
Task 17	After adjusting the CT image to overlay appropriately for target L4, check the registration result using the preview button and select whether to approve it.
Task 18	After selecting whether to approve for target L5, perform image registration again so that the ROI includes all the vertebra area.
Revision of surgical planning	
Task 19	Check the plan on 2D and 3D images, respectively.
Task 20	As a result of planning for the entire target, check whether the robot can move in an area.
Navigation	
Task 21	Insert the screw of target L4-L.
Task 22	Through the [PRE-OP] screen, indicate the values for the insertion depth of the L4-L tapper, the amount of force applied to the end effector, and the patient’s movement.
Task 23	Through the [INTRA-OP] screen, indicate the values for the insertion depth of the L4-L screw, the amount of force applied to the end effector, and the patient’s movement.
Task 24	Move the end effector to the ready position for screw insertion to the target L4-R.

XSL•FO
RenderX

[a]CT: computed tomography.
[b]MPR: multiplanar reconstruction.
[c]AP: anterior-posterior.
[d]OTS: optical tracking system
[e]ROI: region of interest.

XSL•FO
RenderX

Table 2. Use scenarios of nurses.

Use scenario/task number	Task description
Preparation for use	
Task 1	Check the contents related to the emergency button, foot switch, and foot jamming in the user manual.
Task 2	Check the hand jamming label on the robotic arm manipulator.
Task 3	Check if there are any abnormalities in the exterior of the robot marker frame and robotic arm.
Task 4	Place the main console and staff console in a convenient location during surgery.
Task 5	After checking the device and accessories in the operating room, assemble the marker ball.
Task 6	Assemble the surgical tools such as the tapper's driver and marker.
Task 7	Assemble the surgical tools such as screwdriver and marker.
Task 8	Assemble the clamp to be used to fix the patient marker.
System operation	
Task 9	Connect the power and cables of the robotic arm, main console, and staff console.
Task 10	After connecting the foot switch of the robotic arm, turn on the power of the robotic arm.
Initializing manipulator	
Task 11	After logging in, select the surgical method and imaging device.
Task 12	Check the connection status of the foot switch.
Task 13	After selecting the robot position as "right," initialize the manipulator (required to check movement notification sound and operation LED[a]).
Task 14	Verify that the line laser on the robot marker intersects the area within range (required to check movement notification sound and operation LED).
Drape	
Task 15	Follow the on-screen instructions to drape the patient to prevent infection (proceed in order of manipulator drape, base drape, and robot marker drape).
Task 16	After installing the end effector of the robotic arm, assemble the marker ball where the robot marker drape is installed.
Task 17	Move the manipulator to the ready position.
Preparation for surgery	
Task 18	Move the robotic arm for patient surgery.
Task 19	Check the surgical tools through OTS[b], and if all surgical tools are not checked by the OTS camera, check if they are within the operating area.
Task 20	The robot marker is not being recognized by the OTS camera due to damage to the marker ball. Replace with a new marker ball.
Task 21	Please load the surgical data.
Scan	
Task 22	Check if the ROI[c] includes the calibration marker, and the position and direction of the letters "R," "G," and "J" match the image, and then check pass or fail.
Registration	
Task 23	Perform segmentation to distinguish the surgical target in the image. • L4-L, L4-R • L5-L, L5-R
Task 24	Perform labeling to assign target level information of ROI of the 3D image and 2D image segmented for each spine level.
Task 25	Adjust the box so that the ROI of target L4 covers all of the vertebra area.
Task 26	Perform 2D and 3D image registration for each spine level.
Verification	

XSL•FO
RenderX

Use scenario/task number	Task description
Task 27	Use the slide control at the bottom of the image to check whether the 2D and 3D images match to check the registration result.
Task 28	For target L4, move the CT[d] (DRR[e]) image by using the triangular button to adjust the 2 body images to be similar.
Task 29	For target L4 whose registration result has been adjusted, use the preview button to check the registration result and select whether or not to approve it.
Task 30	After selecting whether or not to approve for target L5, perform image registration again so that the ROI includes all of the vertebra area (target L5: registration failed).
Task 31	After displaying the planned data on the screen through the preview button for each target for which the registration result has been adjusted, check if the registration is completed normally.
Task 32	Depending on the registration result of target L5, select whether or not to approve (target L5: registration completed normally).
Planning of surgery	
Task 33	Check whether the robot can move to the planned position.
Navigation	
Task 34	On the screen, move the end effector of the robotic arm to the planned guide position relative to target L4-L.
Task 35	Move the end effector to the original position for the guide.
Task 36	On the screen, move the end effector of the robotic arm to the planned guide position relative to target L4-R.
Use of the emergency stop switch	
Task 37	(At the moment, the manipulator is positioned too close to the patient.) Press the emergency button.
Task 38	Release the emergency button.
Cleaning up after surgery	
Task 39	Shut down the main console.
Task 40	Shut down the robotic arm.
Task 41	Disconnect the cable.

[a]LED: light emitting diode.
[b]OTS: optical tracking system.
[c]ROI: region of interest.
[d]CT: computed tomography.
[e]DRR: digitally reconstructed radiograph.

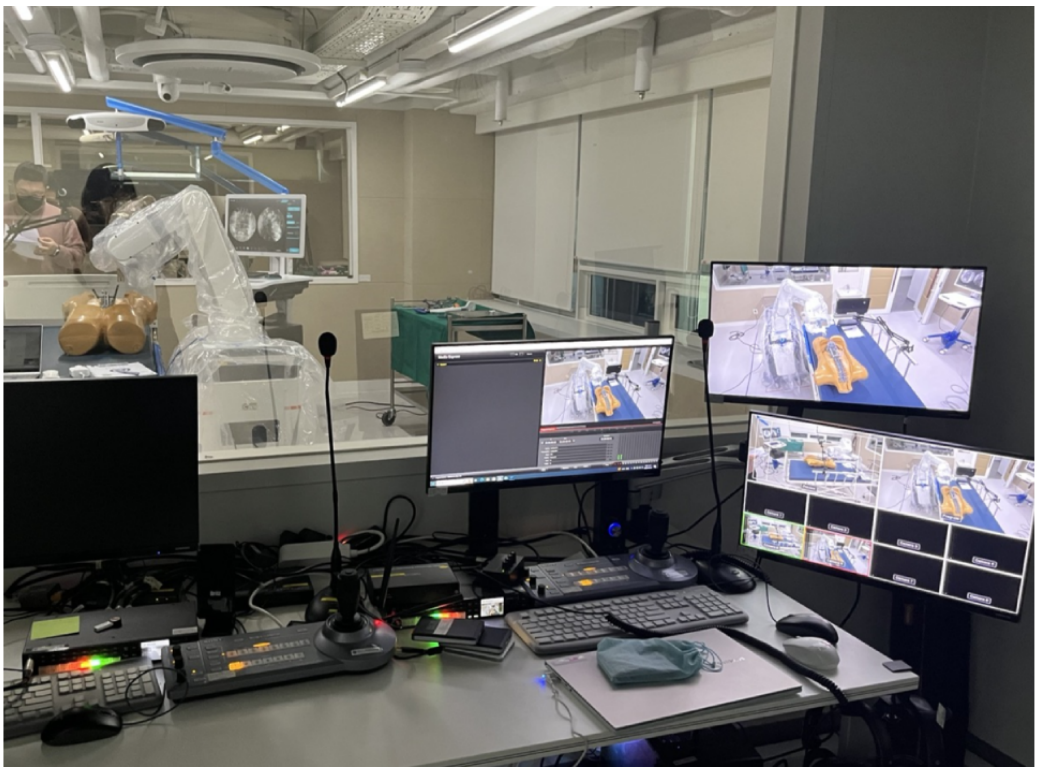
The test environment as shown in Figure 2 is organized to resemble the operating room. Participants used the device following prompts presented on a stand monitor. The test environment was organized similar to an operating room, considering the use environment of the robot spine surgery. An operating room bed, an upper torso dummy, and a patient monitoring device were prepared similar to the actual operating room environment. The temperature and humidity of the evaluation room were measured and recorded right before the evaluation. Similar to a real operating room, the temperature was kept between 20 °C and 24 °C, and the humidity was between 30% and 60%.

The evaluation facilitator guided the participant if they requested assistance with a use scenario, and an observer recorded all participant interactions from outside the test room with a 1-way mirror. The test observation environment setting is shown in Figure 3. The observer used a program from Media Express to record the progress of the usability evaluation. At the end of the evaluation, 3 types of questionnaires were administered.

XSL•FO
RenderX

Figure 2. Test environment: the simulated environment is organized to resemble the operating room in which the evaluator is used.

Figure 3. Test observation environment: we set up monitoring equipment to observe and record the entire evaluation process in real time.



Statistical Analysis

ASQ Measure

After each scenario, the participants completed the ASQ created by Lewis [10] and developed from the ISO 9241-11 standard questionnaire [11]. This is one of the most popular surveys for assessing usability because it is the simplest and its 3 items are easy for participants to understand [11]. As shown in Textbox 1, the ASQ consists of 3 questions, each corresponding to the user's satisfaction with the ease, efficiency according to the time taken to complete the scenario, and validity of the information provided. Participants responded to each question on a 7-point Likert scale [10,12]. Participants rated their satisfaction about the device's usability based on each task scenario [13]. A score of "1" means strongly disagree, and a score of "7" means strongly agree [14]. We found the mean and SD for the 3 questions participants asked ASQ.

XSL•FO
RenderX

Textbox 1. After-Scenario Questionnaire (ASQ).

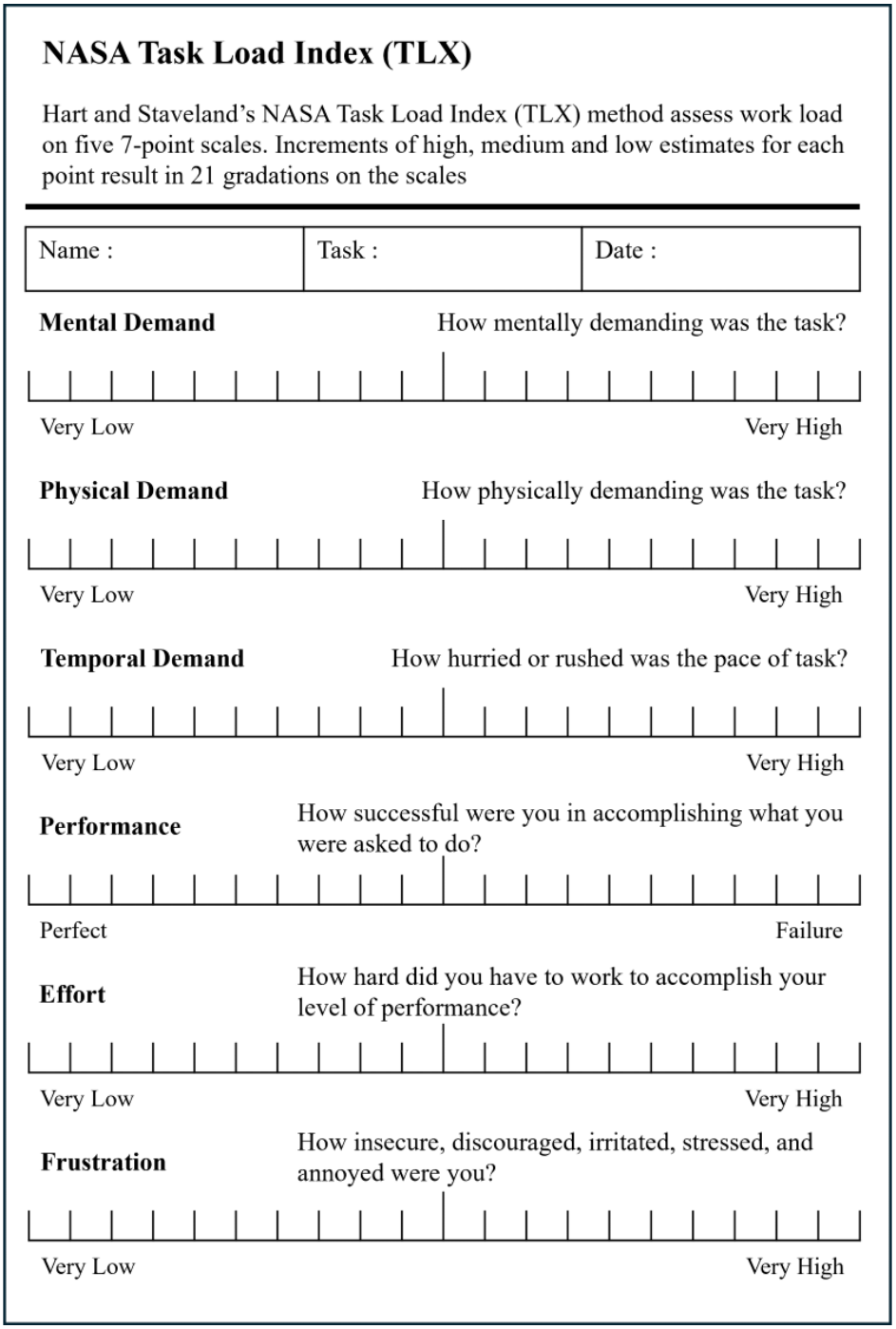
ASQ1: Overall, I am satisfied with the ease of completing the tasks in this scenario.
ASQ2: Overall, I am satisfied with the amount of time it took to complete the tasks in this scenario.
ASQ3: Overall, I am satisfied with the support information (digital help, messages, and documentation) when completing the tasks.

NASA TLX Measure

NASA TLX measures cognitive workload, and like usability, workload is a complex construct that determines the amount of physical and mental effort required to use an interface [15,16]. The workload is assessed by the US NASA TLX [15,17]. The most effective way to assess a worker's perceived job difficulty is to ask questions directly to workers who have experienced the job. As shown in Figure 4, the Task Load Index (TLX) uses 6 dimensions to measure workload. The 6 metrics are mental demand, temporal demand, physical demand, performance, effort, and frustration [16,18]. The NASA TLX scores are evaluated by dividing the score into 21 steps, subtracting 1 from the score, and multiplying it by 5 to express it on a scale of 0 to 100 [19,20]. On a scale of 100, when the score is lower, the workload is lower. Less work means a less complex and easier-to-use user interface. On a 100-point scale, the workload can be described as low (0-9), medium (10-29), rather high (30-49), high (50-79), and very high (80-100) [21]. In the NASA TLX, performance assesses satisfaction with task completion, with the lowest number representing perfect and the highest number representing failure [22]. The point system for mental, physical, temporal, effort, and frustration part ranges from very low to very high [15,16,19,22,23].

XSL•FO
RenderX

Figure 4. NASA Task Load Index.



SUS Measure

As shown in Textbox 2, the SUS consists of 10 items that assess the participant's level of agreement with the overall usability of the system, with odd-numbered items being positive and even-numbered items being negative [24]. SUS is the most commonly used usability assessment questionnaire [24]. Participants responded to each item on a 5-point Likert scale [13]. The scale ranges from 1=strongly disagree to 5=strongly agree [25,26]. To calculate the SUS score from the points acquired from the 5-point Likert scale, the following subtractions were used. For odd-numbered items, subtract 1 from the user response, and for even-numbered items, subtract the user responses from 5. With this calculation, the value range changes from 0 to 4. The most positive response is 4. The scores from each converted response were multiplied by 2.5 to a total possible point of 100. The SUS is percentage-based and divided into 5 levels: A (>80.3), B (68-80.3), C (68), D (51-68), and F (<51) [24]. A score of 85 is considered very good usability, and a score of 68-84 is considered good usability [25,26].

XSL•FO
RenderX

Textbox 2. System Usability Scale (SUS) items.

SUS1: I think that I would like to use this system frequently.

SUS2: I found the system unnecessarily complex.

SUS3: I thought the system was easy to use.

SUS4: I think that I would need the support of a technical person to be able to use this system.

SUS5: I found the various functions in this system were well integrated.

SUS6: I thought there was too much inconsistency in this system.

SUS7: I would imagine that most people would learn to use this system very quickly.

SUS8: I found the system very cumbersome to use.

SUS9: I felt very confident using the system.

SUS10: I needed to learn a lot of things before I could get going with this system.

Data Analysis

ASQ, NASA TLX, and SUS results were computed using SPSS (version 22; IBM Corp) [27]. Descriptive statistics were performed on for doctor and nurse characteristics. Doctors and nurses were compared on age, gender, work experience, and use experience with similar devices. For the questionnaire items, values were compared between groups using 2-tailed *t* tests for normality and Mann-Whitney *U* tests for nonparametric tests. Figures are presented as the mean and SD, and $P<.05$ was considered significant.

Ethical Considerations

This study was approved by the institutional review board of Yonsei University Health System, Gangnam Severance Hospital (3-2022-0493). All the participants who passed the screening signed an informed consent form. Furthermore, all information collected about the participants was anonymized. This study complied with the Code of Ethics. Participants received monetary compensation for participating in the evaluation.

Results

User Statistics

In total, 15 doctors and 15 nurses, each representing the intended users of the surgical assistant, participated in the evaluation. Participants were recruited from doctors and nurses at Severance Hospital. Table 3 shows the sociodemographic characteristics of participants. Both doctors and nurses were between the ages of 30 and 39 years. For both doctors and nurses, those with different experience levels were recruited, and opinions were collected from all the participants. In particular, those with more experience with surgical devices were able to gather relevant opinions because they were more familiar with the device or the existing surgical methods, while those with less experience focused on whether the device was easy to use without much experience. Doctors' professional experience ranged from 2 to 20 years, with an average of 7.53 (SD 5.45) years of professional experience. The nurses' professional experience ranged from 5 to 26 years, with an average of 12.93 (SD 6.43) years. The surgical assistants had used Medtronic (Medtronic), Stryker (Stryker Corp), and Curexo (Curexo, Inc), with an average of 3 (SD 2.36) years of experience.

XSL•FO
RenderX

Table 3. Sociodemographic characteristics and experience of the test participants (N=30).

Variable	Group 1 (doctors), n (%)	Group 2 (nurses), n (%)
Sociodemographic characteristics		
Age (years)		
20-29	3 (20)	1 (7)
30-39	10 (67)	8 (53)
40-49	1 (7)	6 (40)
50-59	1 (7)	0 (0)
Sex		
Male	14 (93)	6 (40)
Female	1 (7)	9 (60)
Work experience		
More than 1 year, less than 5 years	7 (47)	0 (0)
More than 5 years, less than 10 years	3 (20)	4 (27)
More than 10 years, less than 15 years	3 (20)	6 (40)
More than 15 years, less than 20 years	1 (7)	1 (7)
More than 20 years	1 (7)	4 (27)
Use experience with similar devices		
Device name		
Medtronic	12 (80)	7 (47)
Stryker	1 (7)	0 (0)
Curexo	2 (13)	8 (53)
Use experience		
Less than 1 year	6 (40)	3 (20)
More than 1 year, less than 3 years	5 (33)	8 (53)
More than 3 years, less than 5 years	3 (20)	4 (27)
More than 5 years, less than 10 years	0 (0)	0 (0)
More than 10 years	1 (7)	0 (0)

Task Completion

The 15 doctors performed 24 tasks within 8 large scenarios, while the nurses performed a total of 41 tasks within 12 scenarios. As shown in Table 4, for the doctors, all 8 scenarios had a success rate of 90% or higher, with the lowest success rate for the revising a surgical plan scenario. As shown in Table 5, for nurses, all 11 scenarios except the surgical plan had a success rate of 90% or higher, with the planning of surgery scenario having an 87% success rate.

Table 4. Task completion rate in doctors.

	Task pass rate (%)	Task failure rate (%)
Preparations for use	96.67	3.33
Preplanning of surgery	98.67	1.33
Fixation of patient marker	100	0
Scan	96.67	3.33
Registration	100	0
Verification	96.67	3.33
Revision of surgical planning	93.33	6.67
Navigation	98.33	1.67

XSL•FO
RenderX

Table 5. Task completion rate in nurses.

	Task pass rate (%)	Task failure rate (%)
Preparations for use	97.50	2.5
System operations	96.67	3.33
Initialization manipulator	100	0
Drape	93.33	6.67
Preparation for surgery	91.67	8.33
Scan	100	0
Registration	98.33	1.67
Verification	91.11	8.89
Revision of surgical planning	86.67	13.33
Navigation	100	0
Use of the emergency stop switch	100	0
Cleaning up after surgery	95.56	4.44

Among the scenarios in which the doctors did not successfully complete a task during the assessment, the critical tasks were as follows: in task 11, the participant failed to follow the process of attaching the source calibrator in the opposite direction to track the optical tracking system and did not recognize the correct attachment method. In addition, pressing the anterior-posterior scan button and moving the end effector closer to the dummy proceeded correctly, but before attaching the source calibrator, the optical tracking system process could not be performed because it did not proceed in the existing pop-up window and proceeded to the data acquisition step. In task 18, participants selected the target but were unable to click the "Re-matching" button. To proceed with rematching, a target needs to be selected and pressed, but the Re-matching button could not be clicked because the target was not selected. In task 21, the participant did not recognize whether the robot movement was completed by continuously pressing the foot switch without releasing it. In this case, the participant said that he was unable to perform the task because there was no indication on the screen that the robot's movement was complete, and there was no visual or audible user interface.

The nurse was unable to complete task 28 due to difficulty using the image adjustment feature. Participants were asked to move the computed tomography image and adjust the body image to be similar but could not comprehend how to use the "Adjustment" function or the "Re-matching" function (the user did not recognize the intended function itself). Even when the "Adjustment" function was used, it was observed that the user could not use the "Adjustment" function in the way intended by trying to adjust the overlayed screen itself rather than adjusting the screen by pressing the button. If an accurate match is not made, the manipulator may move to a different location than the user's target location, causing potential harm. In addition, nurses had difficulty using the reassembly feature of task 30. The "Re-matching" function could not be used because the target was released while pressing the "Disapproved" button in the rematching task, or the "Adjustment" function was used rather than using the "Re-matching" function. Participants failed to perform the task because they did not recognize that the "Re-matching" function could only be used by resetting the target that was released when pressing the "Disapproved" button, or that "Re-matching" meant rematching. This caused potential harm by moving the manipulator to a location different from the user's target location. Nurses were unsuccessful in tasks such as registration, verification, and navigation because these tasks are usually performed through doctors' orders. During the scenarios, there were no given orders, forcing the nurses to make their own decisions, which they are not accustomed to.

Overall, 4 doctors said that when creating a screw position in the planning stage, the position is created in a completely different part from the actual location; thus, it would be better if the position could be created closer to the target, and when moving the position, that it would be better to be able to check other position paths at the same time. In total, 7 doctors said it would be better if there was notification or guidance for the arrival of the robot arm at the target so that moving to the guided position can be recognized. In addition, 5 doctors and 8 nurses found that in the overall process of selecting and adjusting the region of interest (ROI) box to the target area, it was inconvenient to select and release the box, and that it was difficult to adjust because of its excessive rotation.

Usability (ASQ)

After the usability evaluation, doctors and nurses were surveyed using the ASQ for each scenario. For both doctors and nurses, the ASQ for registration was divided into 2 parts: first, segmentation and labeling, and second, ROI setting and image matching. As shown in Tables 6 and 7, among the registration items, both doctors and nurses had the lowest scores for the ROI setting and image matching, followed by ease of completing the task (group 1: mean 4.73, SD 1.57 and group 2: mean 4.47, SD 1.89), amount of time it took (group 1: mean 4.47, SD 1.63 and group 2: mean 4.40, SD 2.09), and support information satisfaction (group 1: mean 5.13, SD 1.50 and group 2: mean 5.13, SD 1.89). The doctors' opinions were mainly that it was inconvenient to have to click on the line precisely; thus, the ease of adjustment should be improved. Nurses reported that they were less sensitive to the 360-degree rotation button at the

XSL•FO
RenderX

top of the ROI box and had difficulty clarifying the image while adjusting the ROI box.

Table 6. After-Scenario Questionnaire result in group 1 (doctors).

	Ease of completing the task, mean (SD)	Amount of time it took, mean (SD)	Support information satisfaction, mean (SD)
User manual	5.56 (1.12)	5.54 (1.28)	5.63 (1.19)
Preplanning of surgery	5.80 (0.98)	5.53 (1.26)	5.93 (0.77)
Fixation of patient marker	6.00 (0.97)	6.40 (0.61)	__[a]
Scan	6.00 (0.63)	5.60 (0.95)	5.93 (0.77)
Registration 1 (segmentation and labeling)	5.73 (1.18)	5.07 (1.53)	6.00 (0.82)
Registration 2 (ROI[b] setting and image matching)	4.73 (1.57)	4.47 (1.63)	5.13 (1.50)
Verification	5.60 (1.14)	5.40 (1.31)	5.73 (1.00)
Revision of surgical planning	5.80 (0.83)	5.93 (0.68)	5.93 (1.29)
Navigation	5.67 (1.07)	5.93 (0.77)	5.33 (1.45)

[a]Not available; fixation of patient markers was not surveyed because they do not have any on-screen information.

[b]ROI: region of interest.

Table 7. After-Scenario Questionnaire result in group 2 (nurses).

	Ease of completing the task, mean (SD)	Amount of time it took, mean (SD)	Support information satisfaction, mean (SD)
User manual	6.00 (1.23)	5.80 (1.63)	5.97 (1.23)
Preparations for use	6.47 (0.62)	6.07 (1.12)	6.40 (1.02)
System operations	6.27 (0.85)	6.20 (0.98)	6.2 (1.11)
Initialization manipulator	6.13 (1.15)	5.87 (1.45)	5.87 (1.45)
Drape	5.27 (1.84)	5.67 (1.62)	5.80 (1.51)
Preparation for surgery	6.13 (0.88)	6.07 (0.93)	6.27 (0.93)
Scan	6.07 (0.85)	5.87 (0.88)	6.07 (1.06)
Registration 1 (segmentation and labeling)	5.53 (1.50)	4.93 (1.81)	5.67 (1.85)
Registration 2 (ROI[a] setting and image matching)	4.47 (1.89)	4.40 (2.09)	5.13 (1.89)
Verification	4.93 (1.77)	5.20 (1.38)	5.40 (1.74)
Navigation	6.60 (0.61)	6.60 (0.61)	6.60 (0.61)
Use of the emergency stop switch	6.73 (0.44)	6.60 (0.61)	6.73 (0.44)
Cleaning up after surgery	6.47 (0.88)	6.40 (1.02)	6.40 (1.02)

[a]ROI: region of interest.

Tables S1 and S2 in Multimedia Appendix 1 show categorization by use experience with similar devices. The ASQ results did not show significant differences in satisfaction based on use experience and years of experience with robotic surgical systems. For doctors, those with more than 3 years of experience using robotic surgical systems found it easier and faster to perform tasks. For nurses, participants with more experience using similar devices scored higher than those with less than 3 years of experience on the need to prepare before surgery.

Workload (NASA TLX)

Table 8 shows the results of the workload of the assistive robotic surgery devices by occupational group for mental demand, physical demand, temporal demand, performance, effort, frustration, and overall TLX. The Mann-Whitney *U* test comparing the TLX scores of the doctors and nurses, including mental demand (P=.81), physical demand (P=.90), temporal demand (P=.87), performance (P=.81), and frustration (P=.81) and the independent 2-sample *t* test comparing the TLX scores of effort (P=.64) and overall TLX (P=.77) showed no significant differences in the scores. Doctors' workload levels were

XSL•FO
RenderX

generally in the medium (10-29), and nurses were also in the medium (10-29) except for effort.

Table 8. Result of NASA Task Load Index.

	Group 1[a] (n=15), mean (SD)	Group 2[b] (n=15), mean (SD)	*t* test[c] (*df*)	*U* test[d]	*P* value[e]
Mental demand	30.67 (25.97)	29.33 (29.39)	N/A[f]	106.5	.81
Physical demand	13.00 (15.09)	11.33 (12.17)	N/A	109.5	.90
Temporal demand	15.00 (17.63)	22.33 (30.58)	N/A	108.5	.87
Performance	25.00 (28.09)	28.33 (26.16)	N/A	106	.81
Effort	29.67 (24.60)	34.33 (28.78)	–0.477 (28)	N/A	.64
Frustration	19.00 (22.22)	18.33 (26.70)	N/A	106	.81
Overall Task Load Index	22.06 (18.58)	24.00 (16.96)	–0.299 (28)	N/A	.77

[a]Group 1: doctors.

[b]Group 2: nurses.

[c]Because the data were normally distributed, a independent 2-sample *t* test was used.

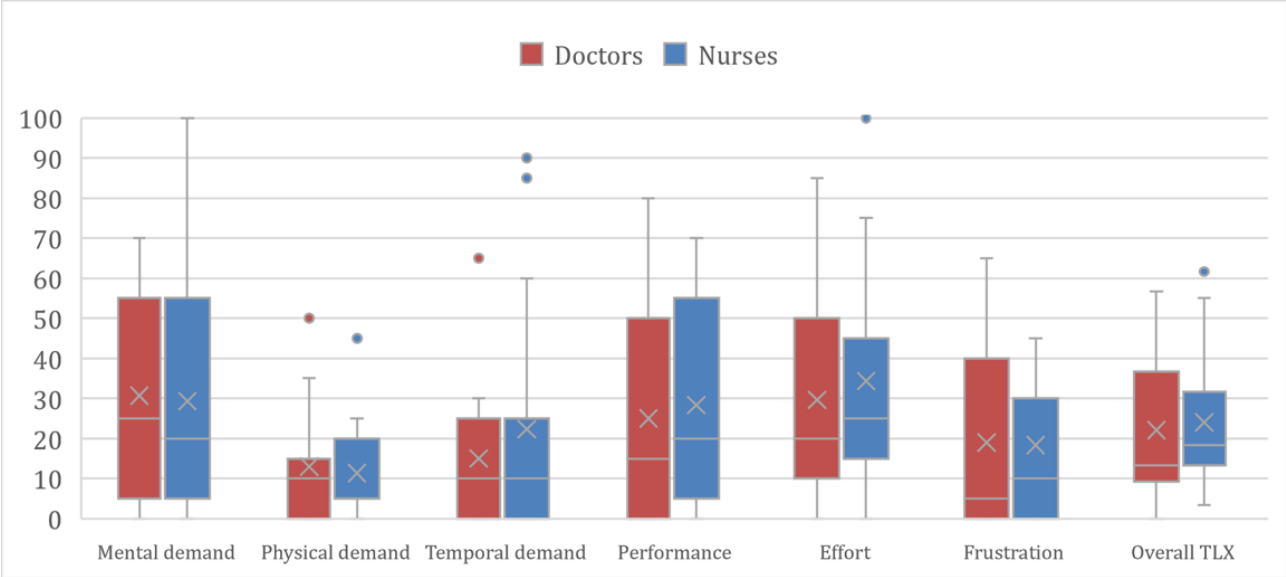
[d]Because the data were not normally distributed, the Mann-Whitney *U* test was performed.

[e]*P* values were determined with the independent 2-sample *t* test and Mann-Whitney *U* test for continuous variables.

[f]N/A: not applicable.

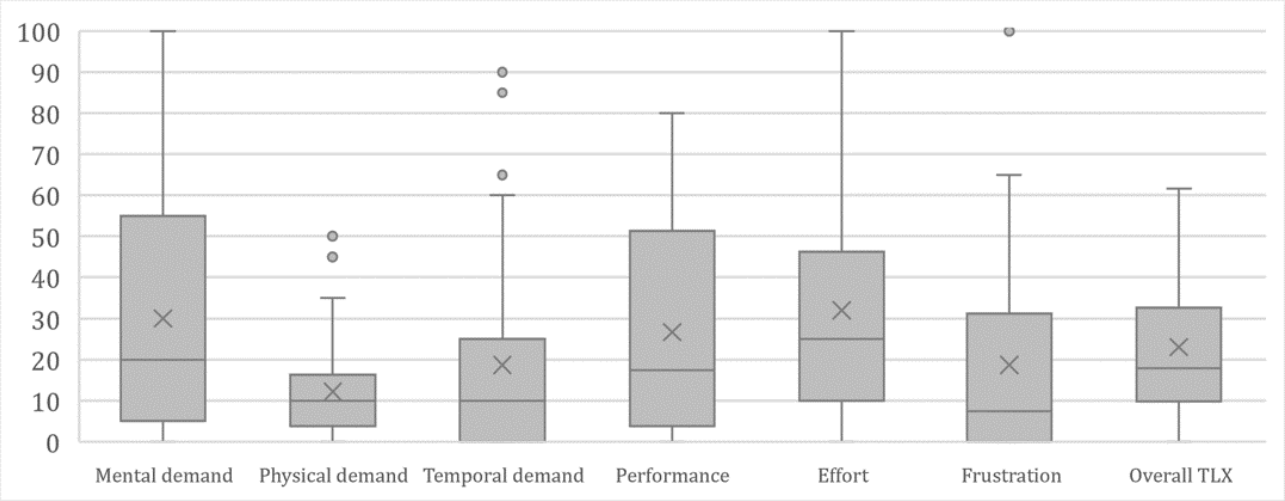
Figure 5 shows a boxplot of the NASA TLX results for doctors and nurses. In addition, Figure 6 shows the NASA TLX results for all evaluation participants (doctors and nurses). The box plots show the maximum (45-100), median (5-25) minimum (0), first quartile (0-17.5), and third quartile (12.5-52.5), with the center box showing the median of 50% of the cases. When comparing the workload of the doctors and nurses, there was no significant difference as shown in Table 8, and we can see that 3 categories, physical demand, temporal demand, and frustration, have lower workloads than the others.

Figure 5. Workload results by group: distribution of the NASA TLX scores for doctors and nurses. TLX: Task Load Index.



XSL•FO
RenderX

Figure 6. Workload results: distribution of the overall NASA TLX scores. TLX: Task Load Index.



In Tables 9 and 10, NASA TLX scores were compared based on use experience with similar devices. Participants who had been using robotic surgical systems for more than 3 years, both doctors and nurses, reported that the evaluation device required a lot of effort to use. Participants had difficulty using the evaluation device because it was more complex than similar robotic surgical devices. However, when it comes to temporal demand, those who have been using robotic surgical systems for more than 3 years reported that it is not time-consuming. The workflow of the evaluation device is not much different from existing robotic surgical devices, and the graphical user interface is easy to use and can be performed quickly.

Table 9. Comparison of the NASA Task Load Index by use experience of doctors.

	All, mean (SD)	Less than 3 years, mean (SD)	More than 3 years, mean (SD)
Mental demand	30.67 (25.97)	27.00 (28.21)	38.00 (21.68)
Physical demand	13.00 (15.09)	14.00 (16.12)	11.00 (14.32)
Temporal demand	15.00 (17.63)	18.50 (19.87)	8.00 (10.37)
Performance	25.00 (28.09)	22.50 (28.80)	30.00 (29.15)
Effort	29.67 (24.60)	21.50 (19.59)	46.00 (27.48)
Frustration	19.00 (22.22)	15.50 (20.20)	26.00 (26.79)

Table 10. Comparison of the NASA Task Load Index by use experience of nurses.

	All, mean (SD)	Less than 3 years, mean (SD)	More than 3 years, mean (SD)
Mental demand	29.33 (29.39)	25.91 (22.56)	38.75 (46.61)
Physical demand	11.33 (12.17)	12.27 (12.72)	8.75 (11.81)
Temporal demand	22.33 (30.58)	28.64 (33.70)	5.00 (5.77)
Performance	28.33 (26.16)	22.27 (24.73)	45.00 (25.50)
Effort	34.33 (28.78)	30.45 (25.73)	45.00 (38.08)
Frustration	18.33 (26.70)	15.45 (15.40)	26.25 (49.22)

Usability (SUS)

Both doctors and nurses who participated in the usability test completed the SUS questionnaire. Table 11 shows that the mean score of SUS was 64.17 (SD 16.52) for doctors and 61.67 (SD 19.18) for nurses. The nonparametric Mann-Whitney *U* test was used to analyze *U* values and *P* values presented in Table 11. When comparing the SUS scores of the doctors and nurses (P=.84, greater than .05), we can see that there was no significant difference between the 2 values. Figure 5 is a boxplot comparing the SUS scores of the doctors and nurses, showing a baseline of 68, which is the average SUS score. For the nurses and doctors, this corresponds to a grade of D on the SUS scale.

XSL•FO
RenderX

Table 11. Result of the System Usability Scale (SUS).

	Group 1[a], mean (SD)	Group 2[b], mean (SD)	*U* test[c]	*P* value[d]
SUS1: I think that I would like to use this system frequently.	2.33 (1.14)	2.60 (1.14)	96.5	.51
SUS2: I found the system unnecessarily complex.	2.4 (1.02)	2.67 (1.25)	92.5	.41
SUS3: I thought the system was easy to use.	2.93 (0.77)	2.93 (1.00)	104	.74
SUS4: I think that I would need the support of a technical person to be able to use this system.	2.00 (1.32)	1.00 (0.97)	65	.05
SUS5: I found the various functions in this system were well integrated.	3.00 (0.82)	2.93 (1.06)	109.5	.90
SUS6: I thought there was too much inconsistency in this system.	2.67 (0.94)	2.47 (1.31)	111	.97
SUS7: I would imagine that most people would learn to use this system very quickly.	2.87 (0.88)	3.20 (1.11)	82	.22
SUS8: I found the system very cumbersome to use.	2.53 (1.09)	2.73 (1.06)	100.5	.62
SUS9: I felt very confident using the system.	2.87 (0.81)	2.73 (0.88)	65	.05
SUS10: I needed to learn a lot of things before I could get going with this system.	2.07 (1.18)	2.00 (1.10)	106.5	.81
Overall, SUS score on 0 to 100 normalized scale	64.17 (16.52)	61.67 (19.18)	107.5	.84

[a]Group 1: doctors.

[b]Group 2: nurses.

[c]Because data were not normally distributed, the Mann-Whitney *U* test was performed.

[d]*P* values were determined with an independent 2-sample *t* test and the Mann-Whitney *U* test for continuous variables.

When we compared participants' use experience with the device between those who had used it for more than 3 years and those who had used it for less than 3 years, we found that for doctors, those who had used it for more than 3 years had lower SUS scores than those who had used it for less than 3 years. For doctors, the average SUS score for participants with 3 or more years of experience is 55.5 (SD 19.13), while the average SUS score for those with less than 3 years is 68.5 (SD 13.05). For nurses, similar to doctors, we found that participants who had used a similar device for more than 3 years had lower scores than those who had used it for less than 3 years, at 53.75 versus 64.55 (Table 12).

Table 12. System Usability Scale (SUS) comparison by use experience of a similar device.

	Group 1[a]			Group 2[b]		
	All, mean (SD)	Less than 3 years, mean (SD)	More than 3 years, mean (SD)	All, mean (SD)	Less than 3 years, mean (SD)	More than 3 years, mean (SD)
SUS	64.17 (16.52)	68.5 (13.05)	55.5 (19.13)	61.67 (19.18)	64.55 (15.14)	53.75 (25.77)

[a]Group 1: doctors.

[b]Group 2: nurses.

Discussion

Principal Findings

This study is a summative evaluation to examine the usability of the frameless stereotaxic navigation system (model CS200) that is used as an auxiliary tool for guiding the surgical tool to the target position and posture planned by the user in the incision or percutaneous spinal surgery. Regarding the use scenario, (1) task success, (2) use error, (3) satisfaction (ASQ), (4) workload (NASA TLX), and (5) SUS related to the usability of the test device were evaluated and analyzed. The usability test was conducted by professional medical staff who have completed specialized medical education and obtained professional medical qualifications. The participants in the usability test were doctors and nurses who have experience in using spinal surgery robots or navigation systems in operating rooms.

For doctors, all 8 scenarios had a success rate of at least 93% or higher, and for nurses, all 12 scenarios had a success rate of at least 87% or higher. Doctors had the lowest success rate of 93% on the "revision of surgical planning" scenario. In the ASQ results, the average score for "ease of completing the task" was 5.66 (SD 0.36), the average score for "amount of time it took" was 5.54 (SD 0.52), and the average score for "support information satisfaction" was 5.70 (SD 0.30). When performing the usability evaluation, the "revision of surgical planning" scenario had the lowest success rate; however, the 3 ASQ scores were higher than the average: 5.80 (SD 0.83), 5.93 (SD 0.68), and 5.93 (SD 1.29).

XSL•FO
RenderX

The "image matching" scenario had the lowest score for each item in the satisfaction score, even though it had a 100% success rate. We found that a high success rate on the evaluation task does not necessarily indicate high usability satisfaction. Despite the high task success rate in the usability test, the low satisfaction rate in the questionnaire that evaluated the usability aspects of the device indicates a lack of satisfaction with the device.

Although there were no difficulties in performance, the task of adjusting the ROI within the "registration" scenario was criticized for its difficulty in accurately adjusting the ROI and its lack of usability. Nurses, like doctors, had the lowest success rate of 87% in the "revision of surgical planning" scenario. However, the ASQ survey results for the "registration" scenario, which had the highest success rate of 98%, showed the lowest scores for "ease of completing the task" with a mean of 4.47 (SD 1.89), "amount of time it took" with a mean of 4.40 (SD 2.09), and "support information satisfaction" with a mean of 5.13 (SD 1.89). Similar to the doctors, when adjusting the ROI box, many of the nurses commented that the 360-degree rotation button at the top was not sensitive, and the video was difficult to see clearly. In addition, both doctors and nurses reported that when moving the robot arm using the footswitch during the "navigation" scenario, there was no visual or audible indication of how far the robot arm had moved and whether it had completed its movement, resulting in collisions between the robot arm and the patient stack. If used with real patients, this could lead to a significant risk of patient injury. We believe that the usability of these screens needs to be improved.

When comparing the workloads of the 2 groups who primarily use the assessment tool, we found that there was no difference, and that the workloads of mental demand and performance are high for both groups. Doctors and nurses commonly commented that the process was too complicated and laborious and that they had doubts about the accuracy of the ROI adjustment. In addition, since the robot assists in surgery, we thought it would be a quick process, but we found that the robot needed more time to move than expected, which affected the workload.

The SUS also showed no significant difference between doctors and nurses, with slightly lower-than-average satisfaction scores. The doctor gave the system a low score on the usability scale because it was too time-consuming to use in the actual operating room. Nurses gave low scores because of the time-consuming setup prior to actual surgery. Both doctors and nurses gave low scores, especially on the items that they felt they would need technician support to use the system and that they would need to learn about the system before its use, because many of them had never used an assessment device before and were not familiar with it. In addition, the lack of usability, with no explanation of what to do next on the device and no prompts to prevent errors, contributed to the low scores. At the hospital where participants work, engineers who have no difficulty using similar devices are present to aid doctors and nurses in using the device. Because of their reliance on engineers, many of the participants had difficulty in using the device alone and commented that they needed the engineers' support. This suggests that the device needs to be highly usable with an easily understandable user interface and a screen design that is familiar to medical staff so that they can use the device without engineers' assistance. Overall, when evaluating usability, there was no significant difference between participants who had used similar devices for more than 3 years and those who had used them for less than 3 years, except for ASQ and SUS, which evaluate satisfaction, and NASA TLX, which evaluates the difficulty of the operator's job. It was found that there were differences in job duties when using the equipment that they were familiar with as well as differences in the time taken to perform the tasks. In other items, there seems to be no problem in using the device once they are familiar with it.

Table S3 in Multimedia Appendix 1 shows the improvements made to the device since the usability evaluation. The user interface was improved by quantifying and intuitively displaying data that used to be shown only as graphs. Confusing highlighted buttons that hindered the use of the device were rearranged to decrease the errors made by users.

To further increase the usability of the device, the footswitch needs to be improved to recognize how much movement is required to move the device by displaying the information on the interface. In particular, during the navigation phase, it would increase the usability if a notification or on-screen guidance appeared when the target was reached while moving to the location guided by the robotic arm. In addition, the drape is divided into 3 stages, while other similar devices only have 1 drape, increasing the risk of contamination.

Although the participants have experience using third-party equipment, they had difficulty using this evaluation device for the first time because they were not familiar with it; however, we do not think there will be any major problems once they are accustomed to the evaluation device. In addition, as a robot that guides the position of the screw in the patient's body, it should have a more accurate and simpler workflow, making it more competitive with other products.

Limitations

This study is limited by the fact that our evaluation took place at a single institution, Severance Hospital in South Korea. In South Korea, robotic surgery is not yet widely used, and many people have not used robotic surgical instruments before; thus, they are still unfamiliar with robotic surgical instruments. However, the strength of this study is that we conducted usability tests with doctors and nurses in the operating room, who are the closest users of the new system, the surgical assistant robot.

Conclusions

In group 1, a success rate of 93% or higher was observed in all 24 tasks. In group 2, a success rate of 87% or higher was observed in 38 of 41 tasks. A success rate of 80% was observed in the task related to marker ball view confirmation (task 18), 80% in the task related to the use of the "Adjustment" function (task 28), and 75% in the task related to using "Re-matching" (task 30).

In addition, subjective data such as follow-up questions and surveys were more effective in identifying shortcomings and judging the usability, satisfaction, and effectiveness of the device

XSL•FO
RenderX

than quantitative data such as the number of use errors (task completion rate) and satisfaction evaluation scores. In terms of error, participants provided a lot of feedback, including suggestions for mitigating potential risks. Although the task success rate was high, the workload and SUS scores were lower than the baseline, suggesting that improving the device user interface would increase the usability of the system. We recommend that the results of this test can be used in other usability engineering processes to improve the overall usability, satisfaction, completeness, and efficiency.

Authors' Contributions

HC wrote the paper and conducted all user experience evaluation, data collection, statistical analysis, and data interpretation. SK conducted user experience evaluation and data collection, such as comparison of the NASA Task Load Index, After-Scenario Questionnaire, and System Usability Scale by use experience. All authors conducted study design and reviewed the final paper.

Conflicts of Interest

None declared.

Multimedia Appendix 1
After-Scenario Questionnaire detailed results for doctors and nurses.
[DOCX File , 3454 KB - humanfactors_v11i1e54425_app1.docx]

References

1. Balicki M, Kyne S, Toporek G, Holthuizen R, Homan R, Popovic A, et al. Design and control of an image-guided robot for spine surgery in a hybrid OR. Int J Med Robot 2020;16(4):e2108. [doi: 10.1002/rcs.2108] [Medline: 32270913]
2. Vadalà G, De Salvatore S, Ambrosio L, Russo F, Papalia R, Denaro V. Robotic spine surgery and augmented reality systems: a state of the art. Neurospine 2020;17(1):88-100 [FREE Full text] [doi: 10.14245/ns.2040060.030] [Medline: 32252158]
3. Lee JW, Kim SM, Kim YS, Chung WK. Automated surgical planning system for spinal fusion surgery with three-dimensional pedicle model. J Inst Control Robot Syst 2011;17(8):807-813 [FREE Full text] [doi: 10.5302/j.icros.2011.17.8.807]
4. Wang TY, Park C, Dalton T, Rajkumar S, McCray E, Owolo E, et al. Robotic navigation in spine surgery: where are we now and where are we going? J Clin Neurosci 2021;94:298-304. [doi: 10.1016/j.jocn.2021.10.034] [Medline: 34863454]
5. Perfetti DC, Kisinde S, Rogers-LaVanne MP, Satin AM, Lieberman IH. Robotic spine surgery: past, present, and future. Spine (Phila Pa 1976) 2022;47(13):909-921. [doi: 10.1097/BRS.0000000000004357] [Medline: 35472043]
6. Mao JZ, Agyei JO, Khan A, Hess RM, Jowdy PK, Mullin JP, et al. Technologic evolution of navigation and robotics in spine surgery: a historical perspective. World Neurosurg 2021;145:159-167. [doi: 10.1016/j.wneu.2020.08.224] [Medline: 32916361]
7. IEC 62366-1:2015. Medical devices: Part 1: Application of usability engineering to medical devices. International Electrotechnical Commission. 2015. URL: https://www.iso.org/standard/63179.html [accessed 2024-03-15]
8. ISO 9241-210:2010. Ergonomics of human-system interaction: Part 210: Human-centred design for interactive systems. International Organization for Standardization. 2010. URL: https://www.iso.org/standard/77520.html [accessed 2024-03-15]
9. Wiklund M, Birmingham L, Larsen SA. Writing Human Factors Plans and Reports for Medical Technology Development. Arlington, VA: Association for the Advancement of Medical Instrumentation; 2018.
10. Lewis JR. Psychometric evaluation of an after-scenario questionnaire for computer usability studies: the ASQ. ACM SIGCHI Bull 1991;23(1):78-81 [FREE Full text] [doi: 10.1145/122672.122692]
11. Muhammad HAB, Hamid K, Iqbal MW, Khurram SK, Shahzad F, Shaheryar M. Usability impact of adaptive culture in smart phones. 2022 Presented at: VIPERC2022: 1st International Virtual Conference on Visual Pattern Extraction and Recognition for Cultural Heritage Understanding; September 12, 2022; Pescara, Italy URL: https://www.researchgate.net/profile/Khalid-Hamid-7/publication/365683592_Usability_Impact_of_Adaptive_Culture_in_Smart_Phones/links/637e4e032f4bca7fd0850d20/Usability-Impact-of-Adaptive-Culture-in-Smart-Phones.pdf
12. Lewis JR. An after-scenario questionnaire for usability studies: psychometric evaluation over three trials. ACM SIGCHI Bull 1991;23(4):79 [FREE Full text] [doi: 10.1145/126729.1056077]
13. Vanterpool KB, Gacki-Smith J, Kuramitsu B, Downey M, Nordstrom MJ, Luken M, et al. A patient-centered website (within reach) to foster informed decision-making about upper extremity vascularized composite allotransplantation: development and usability study. JMIR Form Res 2023;7:e44144 [FREE Full text] [doi: 10.2196/44144] [Medline: 36749618]
14. Yang YF, Chao CJ, Lin YL, Chang CK. Usability testing of Japanese captions segmentation system to scaffold beginners to comprehend Japanese videos. Int J Cyber Soc Educ 2013;6(1):1-14. [doi: 10.7903/ijcse.1083]
15. Hart SG, Staveland LE. Development of NASA-TLX (Task Load Index): results of empirical and theoretical research. Adv Psychol 1988;52:139-183. [doi: 10.1016/s0166-4115(08)62386-9]
16. Hart SG. Nasa-Task Load Index (NASA-TLX); 20 years later. Proc Hum Factors Ergon Soc Annu Meet 2016;50(9):904-908. [doi: 10.1177/154193120605000909]

XSL•FO
RenderX

17. Jiang M, Liu S, Gao J, Feng Q, Zhang Q. A usability study of 3 radiotherapy systems: a comparative evaluation based on expert evaluation and user experience. Med Sci Monit 2019;25:578-589 [FREE Full text] [doi: 10.12659/MSM.913160] [Medline: 30661077]
18. Idrisov B, Hallgren KA, Michaels A, Soth S, Darnton J, Grekin P, et al. Workload, usability, and engagement with a mobile app supporting video observation of methadone take-home dosing: usability study. JMIR Hum Factors 2023;10:e42654 [FREE Full text] [doi: 10.2196/42654] [Medline: 37440298]
19. Hart SG. NASA Task Load Index (TLX). NASA Ames Research Center. 1986. URL: https://ntrs.nasa.gov/citations/20000021487 [accessed 2024-03-12]
20. Patel T, Ivo J, Faisal S, McDougall A, Carducci J, Pritchard S, et al. A prospective study of usability and workload of electronic medication adherence products by older adults, caregivers, and health care providers. J Med Internet Res 2020;22(6):e18073 [FREE Full text] [doi: 10.2196/18073] [Medline: 32348292]
21. Sugarindra M, Suryoputro MR, Permana AI. Mental workload measurement in operator control room using NASA-TLX. IOP Conf Ser Mater Sci Eng 2017;277:012022 [FREE Full text] [doi: 10.1088/1757-899x/277/1/012022]
22. Yurko YY, Scerbo MW, Prabhu AS, Acker CE, Stefanidis D. Higher mental workload is associated with poorer laparoscopic performance as measured by the NASA-TLX tool. Simul Healthc 2010;5(5):267-271 [FREE Full text] [doi: 10.1097/SIH.0b013e3181e3f329] [Medline: 21330808]
23. Kim Y, Son J, Jang W. Usability study on patient monitoring systems: an evaluation of a user interface based on user experience and preference. Med Sci Monit 2023;29:e938570 [FREE Full text] [doi: 10.12659/MSM.938570] [Medline: 37270667]
24. Lewis JR. The system usability scale: past, present, and future. Int J Hum Comput Interact 2018;34(7):577-590. [doi: 10.1080/10447318.2018.1455307]
25. Del Rocio Sevilla-Gonzalez M, Loaeza LM, Lazaro-Carrera LS, Ramirez BB, Rodríguez AV, Peralta-Pedrero ML, et al. Spanish version of the system usability scale for the assessment of electronic tools: development and validation. JMIR Hum Factors 2020;7(4):e21161 [FREE Full text] [doi: 10.2196/21161] [Medline: 33325828]
26. Marzuki MFM, Yaacob NA, Yaacob NM. Translation, cross-cultural adaptation, and validation of the Malay version of the system usability scale questionnaire for the assessment of mobile apps. JMIR Hum Factors 2018;5(2):e10308 [FREE Full text] [doi: 10.2196/10308] [Medline: 29759955]
27. Chaniaud N, Métayer N, Megalakaki O, Loup-Escande E. Effect of prior health knowledge on the usability of two home medical devices: usability study. JMIR Mhealth Uhealth 2020;8(9):e17983 [FREE Full text] [doi: 10.2196/17983] [Medline: 32955454]

Abbreviations

ASQ: After-Scenario Questionnaire
IEC: International Electrotechnical Commission
ISO: International Organization for Standardization
NASA TLX: NASA Task Load Index
ROI: region of interest
SUS: System Usability Scale
TLX: Task Load Index

Edited by A Kushniruk; submitted 27.11.23; peer-reviewed by K Chen, X Cheng; comments to author 17.01.24; revised version received 07.02.24; accepted 26.02.24; published 01.04.24.

Please cite as:
Choi H, Kim S, Jang W
User Experience Evaluation of a Spinal Surgery Robot: Workload, Usability, and Satisfaction Study
JMIR Hum Factors 2024;11:e54425
URL: https://humanfactors.jmir.org/2024/1/e54425
doi:10.2196/54425
PMID:38432688

©Hyeonkyeong Choi, Seunghee Kim, Wonseuk Jang. Originally published in JMIR Human Factors (https://humanfactors.jmir.org), 01.04.2024. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information,

XSL•FO
RenderX

a link to the original publication on https://humanfactors.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Short Paper

Assessing Differences in mHealth Usability and App Experiences Among Young African American Women: Secondary Analysis of a Randomized Controlled Trial

Claudia A Opper[1], BA; Felicia A Browne[1,2], MPH, ScD; Brittni N Howard[1], BA; William A Zule[1], MPH, DrPH; Wendee M Wechsberg[1,2,3], MS, PhD

[1]RTI International, Research Triangle Park, NC, United States

[2]Gillings School of Global Public Health, University of North Carolina, Chapel Hill, NC, United States

[3]Department of Psychiatry and Behavioral Sciences, Duke University School of Medicine, Durham, NC, United States

Corresponding Author:
Felicia A Browne, MPH, ScD
RTI International
3040 East Cornwallis Road
PO Box 12194
Research Triangle Park, NC, 27709-2194
United States
Phone: 1 919 541 6596
Email: fbrowne@rti.org

Abstract

Background: In North Carolina, HIV continues to disproportionately affect young African American women. Although mobile health (mHealth) technology appears to be a tool capable of making public health information more accessible for key populations, previous technology use and social determinants may impact users' mHealth experiences.

Objective: The objective of this study was to evaluate mHealth usability, assessing differences based on previous technology use and social determinants among a sample of African American women in emerging adulthood.

Methods: As part of a National Institute on Drug Abuse–funded randomized controlled trial with African American women (aged 18-25 years), counties were assigned to receive an evidence-based HIV risk reduction intervention through mHealth and participants were asked to complete usability surveys at 6- and 12-month follow-ups. Participants' first survey responses were analyzed through 2-tailed *t* tests and linear regression models to examine associations with previous technology use and social determinants ($P<.05$).

Results: The mean System Usability Scale (SUS) score was 69.2 (SD 17.9; n=159), which was higher than the threshold of acceptability (68.0). Participants who had previously used a tablet indicated higher usability compared to participants without previous use (mean 72.9, SD 18.1 vs mean 57.6, SD 11.4; $P<.001$), and participants with previous smartphone use also reported higher usability compared to participants without previous use (mean 71.9, SD 18.3 vs mean 58.0, SD 10.7; $P<.001$). Differences in SUS scores were observed among those reporting homelessness (mean 58.3, SD 19.0 vs mean 70.8, SD 17.2; $P=.01$), unemployment (mean 65.9, SD 17.2 vs mean 71.6, SD 18.1; $P=.04$), or current school enrollment (mean 73.2, SD 18.5 vs mean 65.4, SD 16.5; $P=.006$). Statistically significant associations were not observed for food insecurity (mean 67.3, SD 18.6 vs mean 69.9, SD 17.7; $P=.45$).

Conclusions: Although above-average usability was observed overall, these findings demonstrate differences in mHealth usability based on past and current life experiences. As mHealth interventions become more prevalent, these findings may have important implications for ensuring that mHealth apps improve the reach of evidence-based interventions.

Trial Registration: ClinicalTrials.gov NCT02965014; https://clinicaltrials.gov/study/NCT02965014

International Registered Report Identifier (IRRID): RR2-10.1186/s12889-018-5796-8

(JMIR Hum Factors 2024;11:e51518) doi:10.2196/51518

KEYWORDS

HIV; Black women; mobile apps; social determinants of health; prevention; substance use; usability

XSL•FO
RenderX

Introduction

In 1981, the first report was published identifying the disease that was later known as AIDS, marking the official beginning of the HIV epidemic [1]. In that same year, IBM's first PC was sold to the public [2]. As the HIV epidemic persists, there may be an opportunity to embrace the digital age we are living in and leverage technological solutions as we work toward the shared goal of ending the HIV epidemic.

HIV incidence rates remain disproportionate based on race in the United States, as rates for Black or African American women are 10.9 times higher than rates for White women [3]. Further, the highest rates of HIV diagnoses occur in the US South [3]. Many interventions have been developed for Black or African American women [4-8], but barriers such as lack of transportation, limited childcare access, concerns over privacy, and community-level stigma impede access to HIV testing and prevention services [9,10].

Most Americans have smart mobile devices [11]. These devices show promise in diminishing barriers and connecting key populations with public health information through increasingly convenient and private pathways. Research has demonstrated that being a woman, young, and African American are characteristics associated with being more likely to prefer mobile health (mHealth) when given a choice or to use mobile devices to seek health information [12,13], supporting mHealth interventions as potentially effective tools for this key population.

However, individuals may experience and engage with mHealth interventions differently. As the prevalence of mHealth apps continues to increase [14], there is a need to understand mHealth usability. In a scoping review of electronic health applications from 2014-2017, the rate of new health applications available outpaced the rate of published usability studies [15]. The authors explained that while most digital health apps are developed commercially, the results of commercial usability studies are not typically published [15]. Because of the limited reported data for mHealth usability, this study examined the usability of an mHealth HIV prevention intervention among young African American women in the US South.

A previous study adapted a best-evidence, women-focused HIV behavioral intervention for young African American women (the Young Women's CoOp), which demonstrated efficacy in reducing sexual risk [16]. In preparation for a trial to test intervention delivery, an mHealth version of the Young Women's Coop was developed [17]. This analysis of usability scores for the Young Women's CoOp mHealth app examines whether previous technology experience and social determinants are associated with mHealth experiences.

Methods

Overview

Analyses in this paper encompass an assessment of the usability of the Young Women's CoOp intervention that was adapted to an mHealth platform. The parent study reached 652 young African American women (aged 18-25 years) in North Carolina who reported recent condomless sex with a male partner and substance use. A complete description of the parent study's eligibility criteria and procedures can be found in the study protocol paper [18].

In a 3-arm randomized trial implementing a cross-over design, 3 counties were assigned to receive the in-person delivery of the Young Women's CoOp intervention, the mHealth delivery of the Young Women's CoOp intervention, or standard HIV counseling and testing. Among the enrolled sample, 197 women were in the counties assigned to receive the mHealth delivery, which consisted of a 1-on-1 mHealth orientation and an Android tablet preloaded with the app that contained the 2-session intervention. Following the orientation, a tablet was provided to each participant to take with them. The study team requested that tablets be returned at the 6-month follow-up appointment.

The usability of the mHealth intervention was evaluated using a modified version of the 10-item System Usability Scale (SUS) [19]. SUS scores range from 0 to 100, with scores above 68 considered above average [20]. To account for participants missing follow-up appointments, mHealth participants completed the usability survey as part of an audio-computer–assisted self-interview (ACASI) at both 6-month and 12-month follow-ups. For participants who completed the usability survey at both follow-ups, only their first (6-month) survey response was considered. ACASI was administered in person in an attempt to engage and collect responses from mHealth users who may have had difficulty using the tablet or lost the tablet and who may have had challenges completing a tablet-hosted survey.

Social determinants (homelessness, unemployment, food insecurity, and school enrollment) were measured at study enrollment. Social determinant variables were either assessed as dichotomous questions (homelessness and school enrollment) or recoded into dichotomous variables (unemployment and food insecurity). Homelessness, unemployment, and school enrollment assessed an individual's current state and food insecurity asked about one's household. Descriptive statistics and 2-tailed t tests were used to assess bivariate associations between social determinants and usability scores. Linear regression was conducted to examine these associations while controlling for previous tablet use and previous smartphone use. Analyses were conducted in Stata 17 (StataCorp) using the threshold of $P<.05$ for statistical significance.

Ethical Considerations

The full study received approval from the Office of Research Protection's Institutional Review Board at RTI International (IRB ID number: 13836). Further, committees from Wake County Human Services and Durham County Department of Public Health, along with administration from the Guilford County Department of Public Health, granted study approval. All participants provided written informed consent. Several procedures were instituted to protect the privacy and confidentiality of study participants, including all staff members involved in data collection and analysis signing and abiding by Staff Agreements of Confidentiality. Additionally, each participant was assigned a unique alphanumeric participant identifier to limit study data being connected to identifying

XSL•FO
RenderX

information, such as name and contact information. Study participants were compensated for completing the baseline appointment with US $50 in gift cards, the 6-month follow-up appointment with US $70 in gift cards, and the 12-month follow-up appointment with US $100 in gift cards.

Results

The overall mean SUS score was 69.2 (SD 17.9; n=159). Less than 12% (n=19) of participants did not have experience with a tablet or smartphone before the study. Variability of SUS scores by previous technology use and social determinants is shown in Table 1.

Participants who had previous tablet use reported higher SUS scores on average than participants who had not previously used a tablet (72.9, SD 18.1 vs 57.6, SD 11.4; *P*<.001). Similarly, participants who had previously used a smartphone had a higher mean SUS score than participants who had not (71.9, SD 18.3 vs 58.0, SD 10.7; *P*<.001).

Additionally, the mean SUS scores were under the acceptable threshold for participants reporting food insecurity, homelessness, unemployment, or no current school enrollment. Statistically significant differences in mean SUS scores were observed among those reporting homelessness (58.3, SD 19.0 vs 70.8, SD 17.2; *P*=.01), unemployment (65.9, SD 17.2 vs 71.6, SD 18.1; *P*=.04), or current school enrollment (73.2, SD 18.5 vs 65.4, SD 16.5; *P*=.006). Statistically significant associations were not observed in the SUS score based on food insecurity (67.3, SD 18.6 vs 69.9, SD 17.7; *P*=.45). When accounting for previous mobile technology experience in each model, homelessness and current school enrollment were statistically significant, but unemployment and food insecurity were not statistically significant (Table 2).

Table 1. Bivariate associations between System Usability Scale (SUS) score and previous technology use and social determinants of health.

	Frequency, n (%)	SUS score, mean (SD)	*P* value
Previous tablet use			<.001
Yes	121 (76.1)	72.9 (18.1)	
No	38 (23.9)	57.6 (11.4)	
Previous smartphone use			<.001
Yes	128 (80.5)	71.9 (18.3)	
No	31 (19.5)	58.0 (10.7)	
Homelessness			.01
Yes	20 (12.6)	58.3 (19.0)	
No	139 (87.4)	70.8 (17.2)	
Unemployment			.04
Yes	67 (42.1)	65.9 (17.2)	
No	92 (57.9)	71.6 (18.1)	
Food insecurity			.45
Yes	41 (25.8)	67.3 (18.6)	
No	118 (74.2)	69.9 (17.7)	
In school			.006
Yes	77 (48.4)	73.2 (18.5)	
No	82 (51.6)	65.4 (16.5)	

Table 2. Associations between System Usability Scale (SUS) score and social determinants of health, adjusting for previous tablet use and smartphone use.

Independent variables	Coefficient (95% CI)	*P* value
Homelessness	–9.0 (–16.8 to –1.1)	.03
Unemployment	–4.9 (–10.1 to 0.3)	.07
Food insecurity	–1.5 (–7.5 to 4.4)	.61
In school	6.2 (1.1 to 11.3)	.02

XSL•FO
RenderX

Discussion

As mHealth continues to become more prevalent, these findings show an overall above-average usability for the mHealth adaptation of the Young Women's CoOp intervention. Notably, there were differences in SUS scores based on a participant's past experiences with technology. Although less than 12% of the study sample had not previously used a smartphone or tablet, this proportion was higher than what may be expected for this age group based on national survey data for young adults [11]. This finding suggests the importance of considering ways the digital divide and other factors may impact familiarity with mobile technology when designing mHealth interventions for this key population. In this study, before participants were given tablets with the mHealth intervention, they received a brief orientation to the app. Given the lower reported usability among those who lacked experience with mobile technology, these findings suggest examining further whether providing more guidance during app orientation could improve app usability for those with limited previous experience. Additionally, ongoing and other technology support (eg, a chat feature where trained staff can provide technology support to users within the app) may be strategies to explore with guidance from the intended end users to see if they improve mHealth experiences for individuals with less mobile technology experience.

Given privacy and stigma-related barriers that may hinder access to in-person HIV prevention programs and services for young African American women [9,10], mHealth could be an attractive solution. However, our findings exemplify that not only previous experience with technology but also diversity in participants' life circumstances, such as homelessness and school enrollment, can be associated with usability. Though the format may appear well-positioned for young adults, it is imperative to consider how a surge in the use of mHealth may miss the opportunity to maximally address existing health disparities if some users encounter barriers when operating the mHealth app that undermine their experiences. Additional guidance and support will be essential for those with factors associated with lower usability.

This study should be considered in relation to a few limitations. All participants completed the intervention using a study-issued Android tablet. Noting how a device's model or operating system may affect usage, some user experiences may have been shaped by the device specifications. In future usability studies, it may be valuable to have participants use their own devices to minimize the chance that device unfamiliarity affects assessments of app usability. Further, participants were asked to assess usability at 6- and 12-month follow-ups. Though participants still had access to the app before returning the device at their follow-up appointment, there was potential for recall bias as usability may have been assessed months after a participant's last app interaction. Additionally, it should be noted that all experiences with the app and data collection occurred before the COVID-19 pandemic. With a greater shift to digital formats for health, education, social, and other services throughout the pandemic, access to mobile devices and familiarity with receiving information through mobile technology may have increased since this study.

Despite these limitations, the study prompts important considerations as the health sector embraces digital technology. The overall above-average usability score signals the potential value of using mHealth as a delivery method in the public health toolkit to further expand the reach of evidence-based interventions to those who may need it the most.

Acknowledgments

The authors would like to thank the participants and the study collaborators at the Durham County Department of Public Health, the Guilford County Department of Public Health, and Wake County Human Services.

This study was funded by the National Institute on Drug Abuse (grant R01 DA041009). The content is solely the work of the authors and does not necessarily reflect the views of the National Institute on Drug Abuse.

Conflicts of Interest

None declared.

References

1. Global HIV/AIDS timeline. KFF. URL: https://www.kff.org/global-health-policy/timeline/global-hivaids-timeline/ [accessed 2022-12-19]
2. Timeline of computer history. Computer History Museum. URL: https://www.computerhistory.org/timeline/1981/ [accessed 2023-01-09]
3. HIV surveillance report, 2020. Vol. 33. Centers for Disease Control and Prevention. 2022. URL: http://www.cdc.gov/hiv/library/reports/hiv-surveillance.html [accessed 2023-07-25]
4. Wechsberg WM, Lam WKK, Zule WA, Bobashev G. Efficacy of a woman-focused intervention to reduce HIV risk and increase self-sufficiency among African American crack abusers. Am J Public Health 2004;94(7):1165-1173 [FREE Full text] [doi: 10.2105/ajph.94.7.1165] [Medline: 15226138]
5. DiClemente RJ, Wingood GM, Harrington KF, Lang DL, Davies SL, Hook EW, et al. Efficacy of an HIV prevention intervention for African American adolescent girls: a randomized controlled trial. JAMA 2004;292(2):171-179 [FREE Full text] [doi: 10.1001/jama.292.2.171] [Medline: 15249566]

XSL•FO
RenderX

6. Gilbert L, Goddard-Eckrich D, Chang M, Hunt T, Wu E, Johnson K, et al. Effectiveness of a culturally tailored HIV and sexually transmitted infection prevention intervention for Black women in community supervision programs: a randomized clinical trial. JAMA Netw Open 2021;4(4):e215226 [FREE Full text] [doi: 10.1001/jamanetworkopen.2021.5226] [Medline: 33835175]
7. Jemmott LS, Jemmott JB, O'Leary A. Effects on sexual risk behavior and STD rate of brief HIV/STD prevention interventions for African American women in primary care settings. Am J Public Health 2007;97(6):1034-1040 [FREE Full text] [doi: 10.2105/AJPH.2003.020271] [Medline: 17463391]
8. Donenberg GR, Kendall AD, Emerson E, Fletcher FE, Bray BC, McCabe K. IMARA: a mother-daughter group randomized controlled trial to reduce sexually transmitted infections in Black/African-American adolescents. PLoS One 2020;15(11):e0239650 [FREE Full text] [doi: 10.1371/journal.pone.0239650] [Medline: 33137103]
9. Mathews A, Farley S, Conserve DF, Knight K, Le'Marus A, Blumberg M, et al. "Meet people where they are": a qualitative study of community barriers and facilitators to HIV testing and HIV self-testing among African Americans in urban and rural areas in North Carolina. BMC Public Health 2020;20(1):494 [FREE Full text] [doi: 10.1186/s12889-020-08582-z] [Medline: 32295568]
10. Rimmler S, Golin C, Coleman J, Welgus H, Shaughnessy S, Taraskiewicz L, et al. Structural barriers to HIV prevention and services: perspectives of African American women in low-income communities. Health Educ Behav 2022;49(6):1022-1032 [FREE Full text] [doi: 10.1177/10901981221109138] [Medline: 35856333]
11. Mobile fact sheet. Pew Research Center. 2021. URL: https://www.pewresearch.org/internet/fact-sheet/mobile/ [accessed 2023-01-15]
12. Khairat S, Liu S, Zaman T, Edson B, Gianforcaro R. Factors determining patients' choice between mobile health and telemedicine: predictive analytics assessment. JMIR mHealth and uHealth 2019;7(6):e13772 [FREE Full text] [doi: 10.2196/13772] [Medline: 31199332]
13. Fox S, Duggan M. Mobile health 2012. Pew Research Center. 2012. URL: https://www.pewinternet.org/wp-content/uploads/sites/9/media/Files/Reports/2012/PIP_MobileHealth2012_FINAL.pdf [accessed 2023-01-09]
14. Bhuyan SS, Lu N, Chandak A, Kim H, Wyant D, Bhatt J, et al. Use of mobile health applications for health-seeking behavior among US adults. J Med Syst 2016;40(6):153. [doi: 10.1007/s10916-016-0492-7] [Medline: 27147516]
15. Maramba I, Chatterjee A, Newman C. Methods of usability testing in the development of eHealth applications: a scoping review. Int J Med Inform 2019;126:95-104. [doi: 10.1016/j.ijmedinf.2019.03.018] [Medline: 31029270]
16. Wechsberg WM, Browne FA, Zule WA, Novak SP, Doherty IA, Kline TL, et al. Efficacy of the Young Women's CoOp: an HIV risk-reduction intervention for substance-using African-American female adolescents in the South. J Child Adolesc Subst Abuse 2017;26(3):205-218 [FREE Full text] [doi: 10.1080/1067828X.2016.1260511] [Medline: 28845096]
17. Watkins RL, Browne FA, Kizakevich PN, Howard BN, Turner LB, Eckhoff R, et al. An evidence-based HIV risk-reduction intervention for young African American women in the US South using mHealth: adaptation and development study. JMIR Form Res 2022;6(5):e34041 [FREE Full text] [doi: 10.2196/34041] [Medline: 35532978]
18. Browne FA, Wechsberg WM, Kizakevich PN, Zule WA, Bonner CP, Madison AN, et al. mHealth versus face-to-face: study protocol for a randomized trial to test a gender-focused intervention for young African American women at risk for HIV in North Carolina. BMC Public Health 2018;18(1):982 [FREE Full text] [doi: 10.1186/s12889-018-5796-8] [Medline: 30081868]
19. System Usability Scale (SUS). US General Services Administration. URL: https://www.usability.gov/how-to-and-tools/methods/system-usability-scale.html [accessed 2022-12-19]
20. Brooke J. SUS: a 'quick and dirty' usability scale. Usability Eval Ind 1996:189-194 [FREE Full text]

Abbreviations

ACASI: audio-computer–assisted self-interview
mHealth: mobile health
SUS: System Usability Scale

Edited by A Kushniruk; submitted 02.08.23; peer-reviewed by D Dunsmuir, C Sun; comments to author 26.11.23; revised version received 17.01.24; accepted 20.02.24; published 16.04.24.

Please cite as:
Opper CA, Browne FA, Howard BN, Zule WA, Wechsberg WM
Assessing Differences in mHealth Usability and App Experiences Among Young African American Women: Secondary Analysis of a Randomized Controlled Trial
JMIR Hum Factors 2024;11:e51518
URL: https://humanfactors.jmir.org/2024/1/e51518
doi:10.2196/51518
PMID:38625721

XSL•FO
RenderX

©Claudia A Opper, Felicia A Browne, Brittni N Howard, William A Zule, Wendee M Wechsberg. Originally published in JMIR Human Factors (https://humanfactors.jmir.org), 16.04.2024. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on https://humanfactors.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Interactive Electronic Pegboard for Enhancing Manual Dexterity and Cognitive Abilities: Instrument Usability Study

Shih-Ying Chien[1,2], PhD; Alice MK Wong[2], MD; Ching-Yi Wu[2,3,4], ScD; Sara L Beckman[5], PhD

[1]Department of Industrial Design, Chang Gung University, Taoyuan, Taiwan
[2]Department of Physical Medicine and Rehabilitation, Chang Gung Medical Foundation, Taoyuan, Taiwan
[3]College of Medicine, Chang Gung University, Taoyuan, Taiwan
[4]Healthy Aging Research Center, Chang Gung University, Taoyuan, Taiwan
[5]Haas School of Business, University of California, Berkeley, Berkeley, CA, United States

Corresponding Author:
Shih-Ying Chien, PhD
Department of Industrial Design
Chang Gung University
No.259, Wen-Hwa 1st Road, Kweishan
Taoyuan, 333
Taiwan
Phone: 886 3 2118800 ext 5669
Email: shihying@mail.cgu.edu.tw

Abstract

Background: Strokes pose a substantial health burden, impacting 1 in 6 people globally. One-tenth of patients will endure a second, often more severe, stroke within a year. Alarmingly, a younger demographic is being affected due to recent lifestyle changes. As fine motor and cognitive issues arise, patient disability as well as the strain on caregivers and health care resources is exacerbated. Contemporary occupational therapy assesses manual dexterity and cognitive functions through object manipulation and pen-and-paper recordings. However, these assessments are typically isolated, which makes it challenging for therapists to comprehensively evaluate specific patient conditions. Furthermore, the reliance on one-on-one training and assessment approaches on manual documentation is inefficient and prone to transcription errors.

Objective: This study examines the feasibility of using an interactive electronic pegboard for stroke rehabilitation in clinical settings.

Methods: A total of 10 patients with a history of stroke and 10 healthy older individuals were recruited. With a limit of 10 minutes, both groups of participants underwent a series of challenges involving tasks related to manual operation, shape recognition, and color discrimination. All participants underwent the Box and Block Test and the Purdue Pegboard Test to assess manual dexterity, as well as an array of cognitive assessments, including the Trail Making Test and the Mini-Mental Status Examination, which served as a basis to quantify participants' attention, executive functioning, and cognitive abilities.

Results: The findings validate the potential application of an interactive electronic pegboard for stroke rehabilitation in clinical contexts. Significant statistical differences ($P<.01$) were observed across all assessed variables, including age, Box and Block Test results, Purdue Pegboard Test outcomes, Trail Making Test-A scores, and Mini-Mental Status Examination performance, between patients with a history of stroke and their healthy older counterparts. Functional and task testing, along with questionnaire interviews, revealed that patients with a history of stroke demonstrated prolonged completion times and slightly inferior performance. Nonetheless, most patients perceived the prototype as user-friendly and engaging. Thus, in the context of patient rehabilitation interventions or the evaluation of patient cognition, physical functioning, or manual dexterity assessments, the developed pegboard could potentially serve as a valuable tool for hand function, attention, and cognitive rehabilitation, thereby mitigating the burden on health care professionals.

Conclusions: Health care professionals can use digital electronic pegboards not only as a precise one-on-one training tool but also as a flexible system that can be configured for online or offline, single-player or multiplayer use. Through data analysis, a more informed examination of patients' cognitive and functional issues can be conducted. Importantly, patient records will be fully retained throughout practices, exercises, or tests, and by leveraging the characteristics of big data, patients can receive the most accurate rehabilitation prescriptions, thereby assisting them in obtaining optimal care.

XSL•FO
RenderX

(JMIR Hum Factors 2024;11:e56357) doi:10.2196/56357

KEYWORDS

interactive electronic pegboard; stroke; hand dexterity; cognitive rehabilitation; system

Introduction

Worldwide, the aging population continues to increase, with several attendant problems [1,2]. The process of aging entails repercussions that extend beyond mere physiological conditions and encompasses a diverse spectrum of complications [3]. Older individuals are confronted with economic, psychological, and societal predicaments stemming from physical aging [4,5]. Hypertension is a critical condition intricately connected with the occurrence of strokes [6,7]. In 2020, a total of 7.08 million individuals globally died due to cerebrovascular disorders [8]. Encouragingly, continuous advancements in medical technology have increased the survival rate for patients with a history of stroke to 62% [9]. Nonetheless, even in cases of survival, 90% of patients experience residual effects, making rehabilitation approaches pivotal [10,11].

Stemming from damage to cerebral tissue, cerebral stroke gives rise to a variety of distinct neurological symptoms contingent upon the site of injury [12,13]. This often culminates in motor, sensory, and cognitive impairments among patients with a history of stroke, in which reduced attentional focus and memory deficits are common [14,15]. The aftermath of a stroke can have negative effects on patients' daily lives, occupational status, and social involvement [16,17]. To enhance physical mobility, manual proficiency, and cognitive aptitude, occupational therapy is the gold standard for elevating overall function [18]. Rehabilitation procedures are initiated once a patient's vital signs stabilize [19]. Clinical evidence indicates that due to significant individual variations among patients with a history of stroke, including age, rehabilitation needs vary [20-22]. Furthermore, older adults predominantly seek to restore ambulatory capacities, whereas younger individuals emphasize intricate fine motor rehabilitation exercises due to occupational demands [23].

In clinical practice, therapists often use calibrated instruments to evaluate and document patients' manual dexterity and cognitive recovery capabilities in one-on-one settings [24-26] using standard methodologies, such as the Purdue Pegboard Test (PPT) and the Nine-Hole Peg Test [24,27]. However, the use of a countdown timer to measure tasks within specific time frames has been validated as an effective means to infer attention, cognition, and manual dexterity capabilities in clinical contexts [28,29]. However, there remain substantial challenges, including human resource depletion, increased time expenditures, difficulties in effective disease progression tracking, and recording errors. Acharya et al [24] emphasized the inherent delay, particularly in response time, with the traditional interactive training method. Compared with the current setup, there is a consistent observation of higher measured timing. Taking the commercially available Neofect Smart Pegboard as an example, it serves as an electronic pegboard [30]. While it offers several advantages, it does not feature long-term tracking and precise prescriptions for individual patients. Furthermore, using the electronic prototype of the Grooved Pegboard Test proposed by Al-Naami et al [31] in 2021 as an example, its operational efficiency shows no significant differences compared with the traditional method of manually recording rehabilitation outcomes. This experiment validates the feasibility of an electronic pegboard test to measure hand-time dexterity with impaired hand functionality, indicating comparable or even superior effectiveness when compared with the conventional manual recording approach.

In this study, we used electronic sensing techniques integrated with Wi-Fi and tablet devices to achieve a higher level of precision in evaluating tasks and time of completion [32]. This digitized approach facilitates accurate documentation of the intricacies associated with each practice and assessment, thereby enhancing the overall precision of the rehabilitation process [33,34]. The principal objective of this study was to subject the prototype to initial evaluation and testing involving patients with a history of stroke and healthy older individuals. We aimed to determine the appropriateness of the set difficulty levels, time constraints, and speed of the prototype.

Methods

Overview

The experimental apparatus consisted of an iPad, 5 color-sensitive building blocks, and 3 variations of task casings. The system's underlying sensing mechanism relied on the modulation of capacitance values resulting from the interaction between the sensing electrodes of the panel and the human body. The conductive building blocks generated stimulation signals that served as surrogate agents for fingers.

A schematic representation is shown in Figure 1. Paired with the distinctive visual patterns on the back of each building block, these visual patterns upon contact with the iPad screen were detected and recognized through pressure sensing. This design was aimed at assessing the responsiveness, visual acuity, and color perception abilities of the participants during rehabilitation interactions (Figure 2).

All rehabilitation tasks and exercises integrated time calculation and countdown functions. Patients were given the option to choose between "independent practice" and "interactive practice" modes. During independent practice, after the "start" button was pressed and the countdown timer initiated, randomized questions were presented. Each practice session was preconfigured for a duration of 10 minutes, and completion and error rates were captured.

For interactive practice, therapists preset practice durations and modify difficulty levels (rehabilitation prescriptions). Through a Wi-Fi connection, therapists administer questions to make an online assessment of patients' abilities. After patients perform the tasks, both patients and therapists receive practice and rehabilitation reports, with all exercise records automatically

XSL•FO
RenderX

stored in the cloud. A flowchart illustrating the operation of the proposed pegboard is shown in Figure 3.

The proposed system encompasses 3 distinct modes, each presenting varying levels of complexity.

Figure 1. Design of the proposed system.

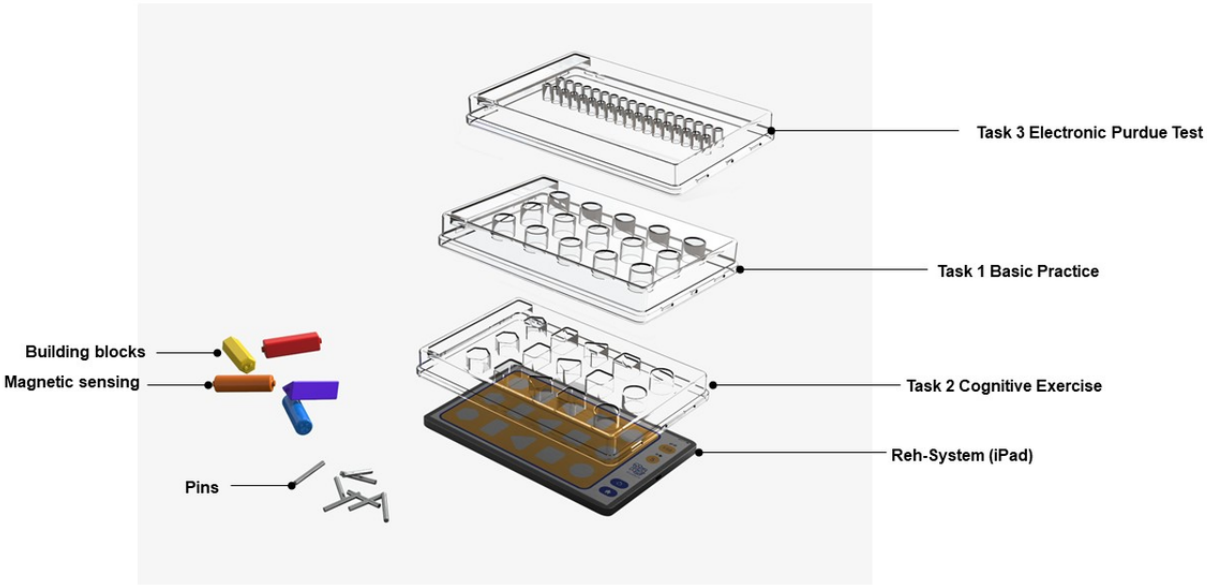
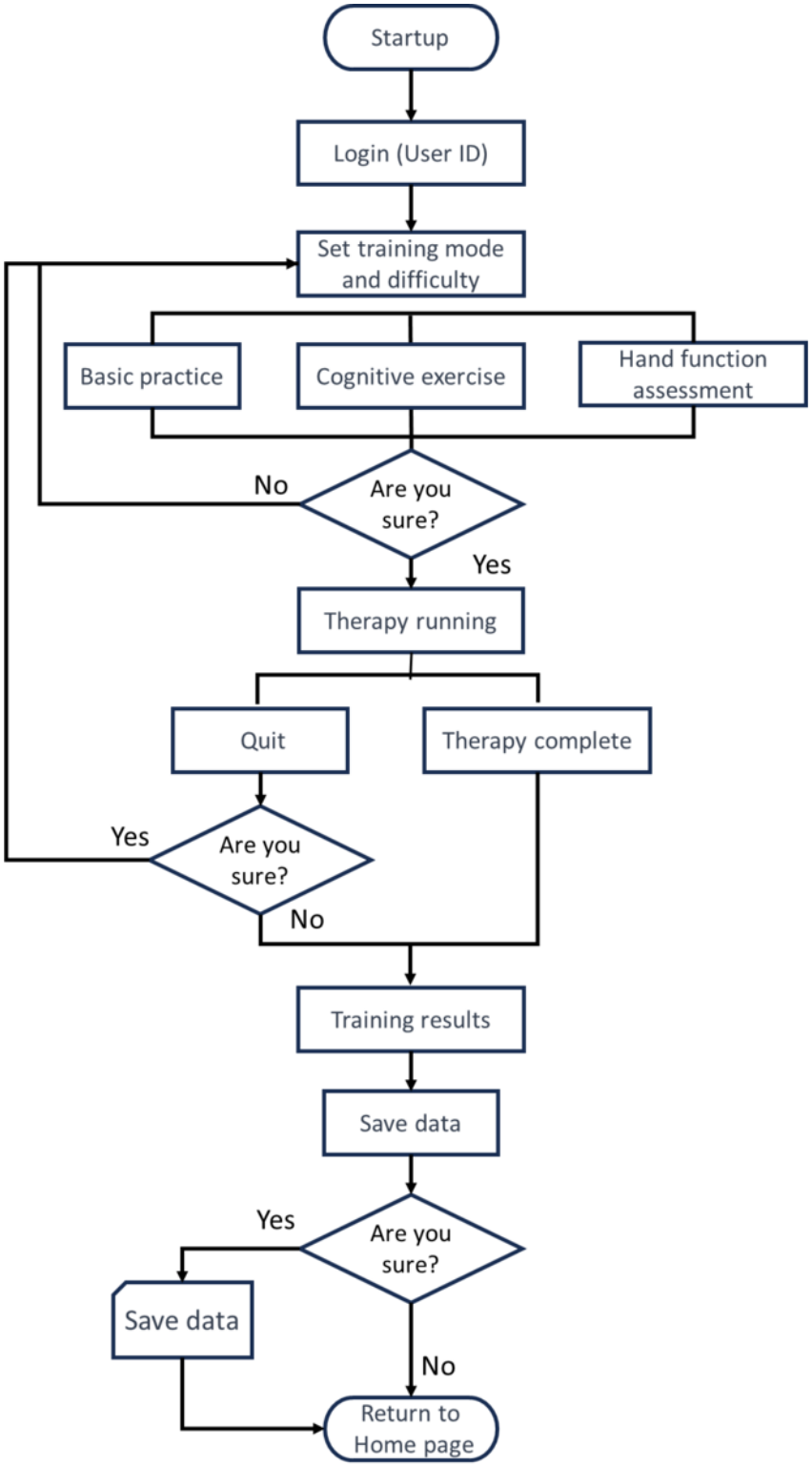


Figure 2. Interactive sensor blocks and tablet interface scenario.

XSL•FO
RenderX

Figure 3. Flowchart of the proposed system operation.



Basic Practice

In basic practice (BP)–1 mode, users are assigned the task of associating the building blocks with the corresponding positions guided by reminder lights on the tablet screen (Figure 4A). Users earn points when they correctly insert the blocks into the panel's corresponding positions. In the intermediate level (BP-2), users place the blocks in the corresponding positions as indicated by the lights of the screen within the given time frame while concurrently considering the variety of colors presented. The advanced level (BP-3) introduces a speed variable to increase the complexity of the task.

XSL•FO
RenderX

Figure 4. Three modes of the proposed system: (A) basic practice; (B) cognitive exercise; and (C) Electronic Purdue Test.

Cognitive Exercise

In cognitive exercise (CE)-1 mode, patients are required to distinguish the shapes of the building blocks and place them according to the patterns displayed on the screen. Users earn points when the blocks are correctly positioned. Upon advancing to the CE-2 level, users need to not only identify the corresponding shaped blocks but also distinguish the colors indicated by the lights. Points are awarded only when both the shape and the color are correct. In addition, this mode introduces varying levels of complexity related to color discrimination and speed (Figure 4B).

Electronic Purdue Test

In the design of the electronic Purdue Test (EPT) level, adherence to the principles of the PPT was paramount. The illuminated signals were meticulously crafted to guide patients in sequentially inserting pegs into corresponding holes (Figure 4C). The assessment consists of distinct 30-second trials for the right, left, and both hands, with individual scores recorded and aggregated. In addition, a 60-second bilateral combination test is administered once. This comprehensive set of evaluations is repeated 3 times. The platform automatically calculates the average score, which serves as the test score. Unlike traditional training, the device guides patients to place pegs into corresponding positions through the use of light signals, which remain illuminated until the pegs are properly placed.

Given the inherent variability in individual patient capabilities, prescribed treatments should differ. To facilitate precision health care, the system incorporates a user login mechanism to generate personalized digital rehabilitation plans and records. The proposed design comprises 3 different modes and 3 difficulty levels of exercises. Preestablished exercises encompass directional movements (including upward, downward, leftward, and rightward motions). During the initial stages of the rehabilitation regimen, the system uses a mechanism of stochastic question generation. As the system accumulates practice data, it systematically discerns and assimilates the individual requirements of each patient, thereby tailoring subsequent questions to enhance areas of observed weakness.

Research Aims

This research had the following aims: (1) we sought to investigate the suitability of the time and difficulty settings for both patients with a history of stroke and healthy users, (2) we explored the correlation between hand function and cognitive abilities, and (3) we conducted a usability questionnaire for the proposed system.

Participant Recruitment

A total of 20 older adults aged between 65 and 80 years were recruited: 10 (50%) patients with a history of stroke and 10 (50%) individuals with no history of strokes. Prior to their inclusion, all participants provided signed informed consent. All participants exhibited right-handed dominance. The inclusion criteria were delineated based on the following: (1) capacity to independently maintain a seated position for a duration exceeding 20 minutes, unaided by external assistance; (2) possession of fundamental communication skills; and (3) relatively uncomplicated functional performance in the assessment of daily life activities. As a safeguard against potential trial-related risks, individuals with a history of recurrent stroke, severe muscular atrophy, or pronounced physical frailty were excluded.

Experimental Procedure

All participants underwent the Box and Block Test (BBT) and the PPT to assess manual dexterity, as well as an array of cognitive assessments, including the Trail Making Test (TMT) and the Mini-Mental Status Examination (MMSE), which served as a basis to quantify participants' attention, executive functioning, and cognitive abilities. Figure 5 shows an image of the proposed system in use.

The experimental session was conducted on a one-on-one basis. The prototype was positioned before each participant, and the 15 pegs, consisting of 5 distinctive colors integral to the interactive system, were methodically arranged adjacent to the central apparatus. Participants sequentially underwent 3 testing modes (BP, CE, and EPT), using their right hand exclusively. Except for the EPT, which followed the PPT criteria, the length of the BP and CE tests was 10 minutes each. The test outcomes encompassed the number of correct responses, completion time, and rehabilitation reports, all of which were concurrently displayed, stored within the apparatus, and uploaded to the cloud platform.

To gain a thorough understanding of users' interactions with the proposed system, we used the System Usability Scale [25] with a 5-point Likert scale to reveal users' perceptions regarding aspects of system acceptance, design appeal, and perceived task difficulty.

XSL•FO
RenderX

Figure 5. Image of the proposed system in use.

Statistical Analysis

This study used SPSS software (version 20; IBM Corp) for statistical analyses. Data analysis involved a comparative assessment between patients with a history of stroke and healthy older individuals, exploring both demographic characteristics and scores obtained from the proposed system, with the Wilcoxon rank sum test used for statistical analysis. To elucidate the potential associations between participants' manual dexterity and cognitive faculties among patients with a history of stroke, this study also used the Spearman rank correlation coefficient, with statistical significance set at $P<.05$.

Ethical Considerations

A total of 20 participants were recruited for this study. All participants provided informed consent by signing a consent form, and the study was conducted in accordance with institutional review board (IRB) regulations using anonymized data, with personal information removed and replaced by codes. No participants withdrew from the study during the research period. To ensure the validity and fairness of the experiment, no monetary or material benefits will be provided during the trial period, in accordance with the IRB application statement. Participants are expected to provide genuine feedback on the product developed in this project based on their intuitive reactions.

The research was conducted within the Department of Physical Medicine and Rehabilitation at Chang Gung Hospital, Taiwan and received approval from the Research Ethics Committee for Human Subject Protection of Chang Gung Medical Foundation (IRB: 202301197A3).

Results

This study included a total of 20 participants (10 patients with a history of stroke and 10 healthy participants). With a limit of 10 minutes, both groups of participants underwent a series of challenges involving tasks related to manual operation, shape recognition, and color discrimination. The statistical analysis revealed statistically significant discrepancies between patients with a history of stroke and healthy participants across all variables ($P<.05$). These differences were evident in all assessed parameters, indicating the potential of the equipment to serve as an assessment tool for both motor and cognitive abilities in both healthy individuals and patients with a history of stroke, with additional training and testing capabilities (Table S1 in Multimedia Appendix 1).

Among older participants who had not experienced a stroke, performance in tasks involving the dominant hand (right hand) during the BBT and the PPT as well as cognitive performance in the TMT was notably superior to those who had experienced a stroke (Figure S1A and S1B in Multimedia Appendix 1). However, no significant differences were observed between the 2 groups in the MMSE test, which assessed memory abilities (Table S2 in Multimedia Appendix 1). Furthermore, in the MMSE test assessing memory abilities, both groups of subjects showed significant differences ($P<.01$) (Table S2 in Multimedia Appendix 1).

XSL•FO
RenderX

Statistical analysis revealed significant negative correlations between performance in the BP-1 or CE-1 task and dexterity tests ($P<.01$). In addition, there were significant correlations between multicolors (BP-2 or CE-2) and dexterity or cognitive tests and a significant negative correlation between scores in the EPT and cognitive performance on the TMT-A and MMSE tests ($P<.05$) (Table S3 in Multimedia Appendix 1). The number of correct answers was used as the score in BP and CE; the time required was used as the score in BP, CE, and EPT in the single and multicolor tests. In terms of usability, 40% (4/10) of patients with a history of stroke and 60% (6/10) of healthy participants deemed the prototype user-friendly (Figure S2A in Multimedia Appendix 1).

Furthermore, all healthy individuals (100%) and majority of patients with a history of stroke (90%, 9/10) found the proposed system highly engaging. During the more demanding CE training, more than 80% (8/10) and 50% (5/10) of healthy participants and patients with a history of stroke, respectively, considered the system both challenging and stimulating. Participants additionally expressed a positive disposition toward the EPT and provided overall positive feedback (Figure S2B in Multimedia Appendix 1).

In assessing task difficulty, nearly 80% (8/10) and 60% (6/10) of healthy participants and patients with a history of stroke, respectively, perceived the BP training tasks as easy and straightforward. However, as participants advanced to the more challenging CT training, there was a noticeable increase in the perceived complexity of tasks, in which only 40% (4/10) and 20% (2/10) of healthy participants and patients with a history of stroke, respectively, found this phase easy. Moreover, 60% (6/10) of healthy participants found the EPT straightforward, while 40% (4/10) of patients with a history of stroke indicated a moderate level of challenge associated with the EPT training (Figure S2C in Multimedia Appendix 1).

Discussion

Principal Findings

Across the 5 tasks investigated in this study, the stroke rehabilitation group exhibited significantly lower scores in the use of the proposed system compared to the healthy participants. In accordance with previous research, advancing age and disease manifest changes and declines in hand function, muscle strength, agility, and cognitive abilities [35,36]. This was evident in the use of the proposed system.

Previous studies have highlighted the repercussions of cerebral damage on patients with a history of stroke, such as compromised cognitive, motor, sensory, and functional capabilities, as well as pain, balance issues, visual challenges, and restricted engagement in activities [37,38]. Sudden cognitive deterioration occurs as a result of these conditions, with 5% of patients with a history of stroke exhibiting dementia symptoms [39,40]. Thus, the augmentation of hand function rehabilitation for patients is imperative. Our findings underscore a close interrelation between manual dexterity and cognitive aptitude [41]. In future research, we intend to explore the use of the proposed system paired with auditory and visual cues to ascertain its potential to guide and enhance visual acuity, attention, and cognitive capabilities among cohorts of different ages and individuals afflicted with cerebral impairments.

The BP and CE tests encompassed factors of both single-color and multicolor conditions. Participants encountered operating difficulties as cognitive demands increased. The CE assessment included recognition of color and object shape, where performance consistently declined across all participants. This observation aligns with previous findings indicating a decline in change detection accuracy with an increase in cognitive load [26,27]. Consequently, as the number of colors increased, the attentional burden on the patients correspondingly increased. This suggests that a graded system with different levels of difficulty might be useful.

During the progression of single-color BP-1 and CE-1 sessions, negative correlations were observed in dexterity tests. In the context of multicolor BP and CE sessions, a distinct and significant correlation was found between the use of multicolors and performance levels on the TMT-A test. The test results suggest that when participants engaged in color recognition and discrimination tasks, the attentional demands for single-color and multicolor tasks differed. In other words, multicolor exercises presented an increased cognitive challenge, affecting manual dexterity and attention switching.

The majority of participants found the proposed system highly user-friendly, in part because the size of the system resembled traditional training pegboards, which maintained familiarity and reduced the need for adaptation. As therapists were not required to manually record participants' actions or time them with stopwatches, this design was advantageous for both users and evaluators. In the various tasks, approximately 60% (12/20) to 70% (14/20) of all participants found BP training to be interesting, while 40% (8/20) to 60% (12/20) of the participants considered the EPT tasks engaging. Regarding the difficulty level, 60% (12/20) to 70% (14/20) of the participants perceived BP training as relatively easy; however, as operating constraints increased (such as color and shape elements), the participants commonly reported that tasks became more challenging and demanding. This finding is consistent with prior research indicating that increasing the difficulty to match users' current abilities enhanced their confidence, maintained attention and engagement in tasks, and promoted a more positive and enjoyable acceptance of new challenges. In this preliminary experiment, neither patients with a history of stroke nor healthy participants were able to complete the tasks within the allotted time, and none of the participants achieved a perfect score. This is likely attributable to the time constraints imposed by the experimental design or the capabilities of the users. Therefore, future studies will include basing task difficulty settings on user performance, similar to leveling up in a video game. The gamification of rehabilitation, in addition to fostering effective interactivity, is facilitated by the incorporation of voice and music assistance. This approach contributes to enhancing the enjoyment of rehabilitation, transforming it from a tedious and uninteresting process. Furthermore, it effectively redirects patients' attention away from pain, thereby augmenting the overall appeal of the rehabilitation process.

XSL•FO
RenderX

The proposed system is equipped for practice, training, and assessment. The majority of similar products on the market predominantly focus on training manual dexterity and do not offer timing and recording functions [28-30]. Certain designs acknowledge the significance of cognitive training and use shape as a cognitive judgment criterion; however, these designs lack elements that enhance the attention of rehabilitation patients, such as auditory cues, visual stimuli, or color-guided prompts. This deficiency in interactive mechanisms often results in users struggling to sustain or commit to rehabilitation efforts. Finally, in terms of assessment functionality, contemporary clinical practice still relies on manual documentation and human intervention for upper limb assessments. The proposed system not only incorporates timing and counting features during upper limb assessments but also introduces guiding and competitive elements, positioning patients to achieve better recovery outcomes. Currently, the system is converting data from each patient's rehabilitation sessions into charts. This aids both patients and health care professionals in gaining a more comprehensive understanding of the rehabilitation and recovery status. This system is thus highly advantageous compared with current commercial products [32-34,42]. The interactive electronic pegboard integrates the merits of existing market offerings and further introduces automated assessment and scoring mechanisms for accurately placed pegs. By surpassing the limitations inherent in conventional fixed training paradigms, this system systematically and comprehensively records the training progress of each case. As data accumulate via a learning model, the system develops a profound understanding of user-specific requirements and consequently extrapolates optimal and customized training regimens tailored to individual users, representing a major step toward precise rehabilitation goals. Finally, the proposed system exhibits greater versatility in its training curriculum and offers increased variability and flexibility. Rehabilitation with digital tools will no doubt significantly enhance users' interest and attention.

Our preliminary investigation indicates that the proposed system is beneficial in the training, assessment, and testing of patients with a history of stroke. The outcomes showcase positive responses concerning hand function training and cognitive ability assessment among patients with a history of stroke. However, to ascertain reliability and validity, a greater number of participants, including diverse age groups and individuals with cerebral impairments, should be recruited in future investigations [35].

Patients with a history of stroke often grapple with diminished motivation for rehabilitation and a lack of immediate feedback, hindering their ability to maintain consistent participation in rehabilitation regimens [36-38]. While the platform devised in this study uses auditory and visual cues to encourage perseverance in rehabilitation, there remains room for improvement in configuring different challenge levels and real-time feedback mechanisms based on varying patient capacities [39-41]. This is a pivotal objective for refinement and enhancement.

Conclusions

The primary objective of this research was to bridge the gap between clinical requirements and product development through customized rehabilitation training based on individual differences. Through the analysis and assessment of data and providing personalized training modes tailored to specific differences, we aimed to predict patients' hand dexterity and cognitive functional abilities. We thus developed the interactive electronic pegboard, a novel software- and hardware-integrated system for stroke rehabilitation, with the purpose of evaluating dexterity and cognitive functions through various task types and multidemonstration patterns. Preliminary findings indicate the efficacy of the system for training and assessment. The ultimate goal of this research is to develop an intelligent system capable of delivering individualized optimized rehabilitation regimens based on the varying needs of users.

Acknowledgments

The authors express their gratitude to all participants in this project for their generous contribution of time, sharing of usage experiences, and insights, which greatly enriched this study. Furthermore, they extend their sincere appreciation to Professor Ching-Yi Wu from the Department of Occupational Therapy at Chang Gung University for providing valuable suggestions and crucial guidance in the experimental design. Finally, they especially thank the Rehabilitation Department at Chang Gung Memorial Hospital for their clinical support throughout this research endeavor. This study received support from the National Science and Technology Council through grant 111-2222-E-182-002-MY2, as well as from the Chang Gung University Innovative Prototype Project with grant EMRPD1M0101.

Data Availability

The data sets generated and analyzed during this study are available from the corresponding author on reasonable request.

Authors' Contributions

SYC was involved in product development and study design, conducted the interviews and prototype testing, collected and analyzed the data, and drafted the manuscript. AMKW and CYW provided clinical guidance. SLB contributed to the manuscript by offering suggestions and participating in revisions.

XSL•FO
RenderX

Conflicts of Interest

None declared.

Multimedia Appendix 1
Detailed participant data and analyses.
[DOCX File , 245 KB - humanfactors_v11i1e56357_app1.docx]

References

1. Bray F, Jemal A, Grey N, Ferlay J, Forman D. Global cancer transitions according to the human development index (2008-2030): a population-based study. Lancet Oncol 2012;13(8):790-801. [doi: 10.1016/S1470-2045(12)70211-5] [Medline: 22658655]
2. Sung H, Ferlay J, Siegel RL, Laversanne M, Soerjomataram I, Jemal A, et al. Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J Clin 2021;71(3):209-249 [FREE Full text] [doi: 10.3322/caac.21660] [Medline: 33538338]
3. McClintock MK, Dale W, Laumann EO, Waite L. Empirical redefinition of comprehensive health and well-being in the older adults of the United States. Proc Natl Acad Sci U S A 2016;113(22):E3071-E3080 [FREE Full text] [doi: 10.1073/pnas.1514968113] [Medline: 27185911]
4. van den Akker M, Buntinx F, Knottnerus JA. Comorbidity or multimorbidity: what's in a name? a review of literature. Eur J Gen Pract 2009;2(2):65-70 [FREE Full text] [doi: 10.3109/13814789609162146]
5. Stein KD, Syrjala KL, Andrykowski MA. Physical and psychological long-term and late effects of cancer. Cancer 2008;112(11 Suppl):2577-2592 [FREE Full text] [doi: 10.1002/cncr.23448] [Medline: 18428205]
6. MacMahon S, Peto R, Cutler J, Collins R, Sorlie P, Neaton J, et al. Blood pressure, stroke, and coronary heart disease. part 1, prolonged differences in blood pressure: prospective observational studies corrected for the regression dilution bias. Lancet 1990;335(8692):765-774. [doi: 10.1016/0140-6736(90)90878-9] [Medline: 1969518]
7. Gorelick PB, Scuteri A, Black SE, Decarli C, Greenberg SM, Iadecola C, et al. Vascular contributions to cognitive impairment and dementia: a statement for healthcare professionals from the American Heart Association/American Stroke Association. Stroke 2011;42(9):2672-2713 [FREE Full text] [doi: 10.1161/STR.0b013e3182299496] [Medline: 21778438]
8. Manji RA, Arora RC, Singal RK, Hiebert BM, Menkis AH. Early rehospitalization after prolonged intensive care unit stay post cardiac surgery: outcomes and modifiable risk factors. J Am Heart Assoc 2017;6(2):e004072 [FREE Full text] [doi: 10.1161/JAHA.116.004072] [Medline: 28174166]
9. Filsoufi F, Rahmanian PB, Castillo JG, Bronster D, Adams DH. Incidence, topography, predictors and long-term survival after stroke in patients undergoing coronary artery bypass grafting. Ann Thorac Surg 2008;85(3):862-870. [doi: 10.1016/j.athoracsur.2007.10.060] [Medline: 18291158]
10. Hoskins WJ, McGuire WP, Brady MF, Homesley HD, Creasman WT, Berman M, et al. The effect of diameter of largest residual disease on survival after primary cytoreductive surgery in patients with suboptimal residual epithelial ovarian carcinoma. Am J Obstet Gynecol 1994;170(4):974-979. [doi: 10.1016/s0002-9378(94)70090-7] [Medline: 8166218]
11. Wasserstrom WR, Glass JP, Posner JB. Diagnosis and treatment of leptomeningeal metastases from solid tumors: experience with 90 patients. Cancer 1982;49(4):759-772 [FREE Full text] [doi: 10.1002/1097-0142(19820215)49:4<759::aid-cncr2820490427>3.0.co;2-7] [Medline: 6895713]
12. Gaetz M. The neurophysiology of brain injury. Clin Neurophysiol 2004;115(1):4-18. [doi: 10.1016/s1388-2457(03)00258-x] [Medline: 14706464]
13. Schmahmann JD, Smith EE, Eichler FS, Filley CM. Cerebral white matter: neuroanatomy, clinical neurology, and neurobehavioral correlates. Ann N Y Acad Sci 2008;1142(1):266-309 [FREE Full text] [doi: 10.1196/annals.1444.017] [Medline: 18990132]
14. Mahncke HW, Bronstone A, Merzenich MM. Brain plasticity and functional losses in the aged: scientific bases for a novel intervention. Prog Brain Res 2006;157:81-109. [doi: 10.1016/S0079-6123(06)57006-2] [Medline: 17046669]
15. Prokopenko SV, Mozheyko EY, Petrova MM, Koryagina TD, Kaskaeva DS, Chernykh TV, et al. Correction of post-stroke cognitive impairments using computer programs. J Neurol Sci 2013;325(1-2):148-153. [doi: 10.1016/j.jns.2012.12.024] [Medline: 23312291]
16. Hertzog C, Kramer AF, Wilson RS, Lindenberger U. Enrichment effects on adult cognitive development: can the functional capacity of older adults be preserved and enhanced? Psychol Sci Public Interest 2008;9(1):1-65 [FREE Full text] [doi: 10.1111/j.1539-6053.2009.01034.x] [Medline: 26162004]
17. Liberman RP. Recovery From Disability: Manual of Psychiatric Rehabilitation. Washington, DC: American Psychiatric Publishing; 2009.
18. Pendleton HM, Schultz-Krohn W. Pedretti's Occupational Therapy: Practice Skills for Physical Dysfunction. St. Louis, MO: Elsevier Health Sciences; 2017.

XSL•FO
RenderX

19. Duncan PW, Zorowitz R, Bates B, Choi JY, Glasberg JJ, Graham GD, et al. Management of adult stroke rehabilitation care: a clinical practice guideline. Stroke 2005;36(9):e100-e143 [FREE Full text] [doi: 10.1161/01.STR.0000180861.54180.FF] [Medline: 16120836]
20. Quinn TJ, Paolucci S, Sunnerhagen KS, Sivenius J, Walker MF, Toni D, et al. Evidence-based stroke rehabilitation: an expanded guidance document from the European Stroke Organisation (ESO) guidelines for management of ischaemic stroke and transient ischaemic attack 2008. J Rehabil Med 2009;41(2):99-111 [FREE Full text] [doi: 10.2340/16501977-0301] [Medline: 19225703]
21. Heinemann AW, Linacre JM, Wright BD, Hamilton BB, Granger C. Relationships between impairment and physical disability as measured by the functional independence measure. Arch Phys Med Rehabil 1993;74(6):566-573. [doi: 10.1016/0003-9993(93)90153-2] [Medline: 8503745]
22. Winstein CJ, Stein J, Arena R, Bates B, Cherney LR, Cramer SC, et al. Guidelines for adult stroke rehabilitation and recovery: a guideline for healthcare professionals from the American Heart Association/American Stroke Association. Stroke 2016;47(6):e98-e169 [FREE Full text] [doi: 10.1161/STR.0000000000000098] [Medline: 27145936]
23. Majmudar S, Wu J, Paganoni S. Rehabilitation in amyotrophic lateral sclerosis: why it matters. Muscle Nerve 2014;50(1):4-13 [FREE Full text] [doi: 10.1002/mus.24202] [Medline: 24510737]
24. Acharya KA, Bhat S, Kanthi M, Rao BK. Fine motor assessment in upper extremity using custom-made electronic pegboard test. J Med Signals Sens 2022;12(1):76-83 [FREE Full text] [doi: 10.4103/jmss.JMSS_58_20] [Medline: 35265469]
25. Bangor A, Kortum PT, Miller JT. An empirical evaluation of the system usability scale. Int J Hum Comput 2008;24(6):574-594 [FREE Full text] [doi: 10.1080/10447310802205776]
26. Vrij A, Mann SA, Fisher RP, Leal S, Milne R, Bull R. Increasing cognitive load to facilitate lie detection: the benefit of recalling an event in reverse order. Law Hum Behav 2008;32(3):253-265. [doi: 10.1007/s10979-007-9103-y] [Medline: 17694424]
27. Setz C, Arnrich B, Schumm J, La Marca R, Tröster G, Ehlert U. Discriminating stress from cognitive load using a wearable EDA device. IEEE Trans Inf Technol Biomed 2010;14(2):410-417 [FREE Full text] [doi: 10.1109/TITB.2009.2036164] [Medline: 19906598]
28. Curcio G, Ferrara M, De Gennaro L. Sleep loss, learning capacity and academic performance. Sleep Med Rev 2006;10(5):323-337. [doi: 10.1016/j.smrv.2005.11.001] [Medline: 16564189]
29. Schmidt RA, Lee TD, Winstein CJ, Wulf G, Zelaznik HN. Motor control and Learning: A Behavioral Emphasis. Champaign, IL: Human Kinetics; 2018.
30. Ericsson KA, Charness N. Expert performance: its structure and acquisition. Am Psychol 1994;49(8):725-747 [FREE Full text] [doi: 10.1037//0003-066x.49.8.725]
31. Al-Naami B, Al-Naimat F, Almalty A, Visconti P, Al-Hinnawi A. A prototype of an electronic pegboard test to measure hand-time dexterity with impaired hand functionality. ASI 2021 Dec 22;5(1):2. [doi: 10.3390/asi5010002]
32. Lee JS, Kim SW. Effects of Korean computer-based cognitive rehabilitation program on the memory in healthy elderly. JIAPTR 2018;9(4):1591-1595 [FREE Full text] [doi: 10.20540/JIAPTR.2018.9.4.1591]
33. Ham Y, Shin JH. Efficiency and usability of a modified pegboard incorporating computerized technology for upper limb rehabilitation in patients with stroke. Top Stroke Rehabil 2023;30(4):333-341 [FREE Full text] [doi: 10.1080/10749357.2022.2058293] [Medline: 35348031]
34. Okahashi S, Sakamoto K, Hashiya F, Kumasaka K, Yamaguchi T, Seiyama A, et al. Development of an electric pegboard (e-Peg) for hand dexterity improvement and cognitive rehabilitation: a preliminary study. Adv Biomed Eng 2023;12:81-90 [FREE Full text] [doi: 10.14326/abe.12.81]
35. Weintraub S, Salmon D, Mercaldo N, Ferris S, Graff-Radford NR, Chui H, et al. The Alzheimer's Disease Centers' Uniform Data Set (UDS): the neuropsychologic test battery. Alzheimer Dis Assoc Disord 2009;23(2):91-101 [FREE Full text] [doi: 10.1097/WAD.0b013e318191c7dd] [Medline: 19474567]
36. Taylor J, Taylor S. Psychological Approaches to Sports Injury Rehabilitation. Gaithersburg, MD: Aspen Publishers; 1997.
37. Sohlberg MM, Mateer CA. Cognitive Rehabilitation: An Integrative Neuropsychological Approach. New York, NY: Guilford Press; 2001.
38. Borghese NA, Pirovano M, Lanzi PL, Wüest S, de Bruin ED. Computational intelligence and game design for effective at-home stroke rehabilitation. Games Health J 2013;2(2):81-88 [FREE Full text] [doi: 10.1089/g4h.2012.0073] [Medline: 24761321]
39. Sung M, Marci C, Pentland A. Wearable feedback systems for rehabilitation. J Neuroeng Rehabil 2005;2:17 [FREE Full text] [doi: 10.1186/1743-0003-2-17] [Medline: 15987514]
40. Sveistrup H. Motor rehabilitation using virtual reality. J Neuroeng Rehabil 2004;1(1):10 [FREE Full text] [doi: 10.1186/1743-0003-1-10] [Medline: 15679945]
41. Novak D. Promoting motivation during robot-assisted rehabilitation. In: Colombo R, Sanguineti V, editors. Rehabilitation Robotics. St. Louis, MO: Elsevier Science; 2018:149-158.
42. Shin JH, Kim MY, Lee JY, Jeon YJ, Kim S, Lee S, et al. Effects of virtual reality-based rehabilitation on distal upper extremity function and health-related quality of life: a single-blinded, randomized controlled trial. J Neuroeng Rehabil 2016;13(1):17 [FREE Full text] [doi: 10.1186/s12984-016-0125-x] [Medline: 26911438]

XSL•FO
RenderX

Abbreviations

BBT: Box and Block Test
BP: basic practice
CE: cognitive exercise
EPT: Electronic Purdue Test
MMSE: Mini-Mental Status Examination
PPT: Purdue Pegboard Test
TMT: Trail Making Test

Edited by Y Gong; submitted 14.01.24; peer-reviewed by YW Shau, C Baxter; comments to author 04.03.24; revised version received 22.03.24; accepted 04.04.24; published 21.06.24.

Please cite as:
Chien SY, Wong AMK, Wu CY, Beckman SL
Interactive Electronic Pegboard for Enhancing Manual Dexterity and Cognitive Abilities: Instrument Usability Study
JMIR Hum Factors 2024;11:e56357
URL: https://humanfactors.jmir.org/2024/1/e56357
doi:10.2196/56357
PMID:38904991

©Shih-Ying Chien, Alice MK Wong, Ching-Yi Wu, Sara L Beckman. Originally published in JMIR Human Factors (https://humanfactors.jmir.org), 21.06.2024. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on https://humanfactors.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Effects of User Experience in Automated Information Processing on Perceived Usefulness of Digital Contact-Tracing Apps: Cross-Sectional Survey Study

Tim Schrills[1], BSc, MSc; Lilian Kojan[1], BSc, MSc; Marthe Gruner[1], BSc, MSc; André Calero Valdez[1], Prof Dr; Thomas Franke[1], Prof Dr

Institute for Multimedia and Interactive Systems, Universität zu Lübeck, Lübeck, Germany

Corresponding Author:
Tim Schrills, BSc, MSc
Institute for Multimedia and Interactive Systems
Universität zu Lübeck
Ratzeburger Allee 160
Lübeck, 23560
Germany
Phone: 49 451 3101 ext 5135
Fax: 49 451 3101 5104
Email: Tim.schrills@uni-luebeck.de

Abstract

Background: In pandemic situations, digital contact tracing (DCT) can be an effective way to assess one's risk of infection and inform others in case of infection. DCT apps can support the information gathering and analysis processes of users aiming to trace contacts. However, users' use intention and use of DCT information may depend on the perceived benefits of contact tracing. While existing research has examined acceptance in DCT, automation-related user experience factors have been overlooked.

Objective: We pursued three goals: (1) to analyze how automation-related user experience (ie, perceived trustworthiness, traceability, and usefulness) relates to user behavior toward a DCT app, (2) to contextualize these effects with health behavior factors (ie, threat appraisal and moral obligation), and (3) to collect qualitative data on user demands for improved DCT communication.

Methods: Survey data were collected from 317 users of a nationwide-distributed DCT app during the COVID-19 pandemic after it had been in app stores for >1 year using a web-based convenience sample. We assessed automation-related user experience. In addition, we assessed threat appraisal and moral obligation regarding DCT use to estimate a partial least squares structural equation model predicting use intention. To provide practical steps to improve the user experience, we surveyed users' needs for improved communication of information via the app and analyzed their responses using thematic analysis.

Results: Data validity and perceived usefulness showed a significant correlation of r=0.38 (P<.001), goal congruity and perceived usefulness correlated at r=0.47 (P<.001), and result diagnosticity and perceived usefulness had a strong correlation of r=0.56 (P<.001). In addition, a correlation of r=0.35 (P<.001) was observed between Subjective Information Processing Awareness and perceived usefulness, suggesting that automation-related changes might influence the perceived utility of DCT. Finally, a moderate positive correlation of r=0.47 (P<.001) was found between perceived usefulness and use intention, highlighting the connection between user experience variables and use intention. Partial least squares structural equation modeling explained 55.6% of the variance in use intention, with the strongest direct predictor being perceived trustworthiness (β=.54; P<.001) followed by moral obligation (β=.22; P<.001). Based on the qualitative data, users mainly demanded more detailed information about contacts (eg, place and time of contact). They also wanted to share information (eg, whether they wore a mask) to improve the accuracy and diagnosticity of risk calculation.

Conclusions: The perceived result diagnosticity of DCT apps is crucial for perceived trustworthiness and use intention. By designing for high diagnosticity for the user, DCT apps could improve their support in the action regulation of users, resulting in higher perceived trustworthiness and use in pandemic situations. In general, automation-related user experience has greater importance for use intention than general health behavior or experience.

(JMIR Hum Factors 2024;11:e53940) doi:10.2196/53940

XSL•FO
RenderX

KEYWORDS

COVID-19; contact tracing; user experience; trust; health information processing

Introduction

Background

During pandemic situations, efficiently acquiring, storing, and evaluating information on physical contacts can be crucial for both individuals and public health agencies aiming to curb infection dynamics [1]. Manual tracing of such contacts is practically impossible, leading to a growing development and research of digital tools supporting such efforts, commonly referred to as digital contact tracing (DCT) apps [2]. By allowing for automation, DCT tools effectively allow for contact tracing. They aim to allow individual users to assess their own risk status with minimal effort and offer support in daily action regulation, such as in decision situations, regarding isolation or notification of previous contacts [3]. If used correctly, DCT can aid in breaking chains of infection and thereby support curbing pandemic spread. For example, in Germany, a DCT called *Corona-Warn-App* (CWA) [4] was developed on behalf of the Federal Ministry of Health, and it was downloaded >40 million times [5].

However, the extent to which individuals use DCT can vary vastly [6]. Previous research has shown that it is crucial whether users perceive a DCT app as beneficial to guide them in pandemic contexts [7]. This core factor is in line with existing models of health behavior (eg, the influential Health Belief Model [HBM] [8]). Within the HBM, perceived benefit is outlined as a central determinant for the implementation of health behavior [7]. When investigating health-related technology, the HBM is frequently connected with models of technology acceptance [9]. As part of these models, the perceived usefulness or performance of technology is similarly postulated as a central variable for use intention. In this paper, we refer to the term *usefulness* as it is better suited than *benefits* to describe the effects of a specific technology. Thereby, we refer to usefulness as "the degree to which a person believes that using a particular system would enhance their [...] performance" [10].

Examining psychological processes revolving around the perception of DCT usefulness is a crucial research topic to understand the adoption and efficient implementation of DCT. Extensive research has shown the importance of the perceived usefulness of DCT for different applications and in different countries [11-15]. All in all, extending existing theoretical approaches such as the HBM by focusing on user experience variables in DCT allows for clear guidelines on improving DCT design and uptake.

The usefulness that a user can experience from DCT results from the automation it provides. DCT takes over tasks that would otherwise need to be done manually (eg, recording contacts, estimating distance and exposure to contacts, and calculating risk based on the vaccination status of contacts). Therefore, it can be defined as an automated system. In general, automation can be defined as a system's ability to "offload, assist, or replace human performance at corresponding stages of human information processing" [16]. The human action that DCT seeks to automate is the continuous recording and analysis of contact data to monitor an individual's risk of infection. While there is a large body of research on automation, its adverse biases, and its impact on human performance [17-19], less research focuses on the psychological processes involved when users evaluate the usefulness of automated contact tracing.

Parasuraman et al [20] define 4 evaluation criteria on how automation can affect human performance: situation awareness [21], trust (cf complacency and trust [22]), skill degradation [23], and workload [24]. When users want to make situation-adequate decisions, they benefit from improved situation awareness. Situation awareness, in turn, can be improved by DCT. As long as the information or recommendations provided by DCT apps are perceived as trustworthy, users may use them to determine the right course of action. Accordingly, a DCT's ability to support situation awareness as well as trust formation (refer to the study by Hoff and Bashir [25]) may lead to perceived usefulness. On the other hand, in the context of DCT apps, one cannot assume that users are potentially losing a previously existing skill through automation; DCT app users are not able to stop sick individuals or themselves. Along the same line, DCT app users profit from automation as it reduces manual work in contact tracing. Therefore, we propose to examine users' experience of situation awareness and trustworthiness when using DCT apps.

While research has demonstrated that usefulness strongly impacts use intention [26], factors unrelated to the specific DCT app might affect whether people intend to use the system. The HBM positions threat appraisal as another factor directly influencing use intention [7]. While using a DCT app changes neither the susceptibility nor the severity (in comparison, refer to the study by Costa [27]) related to an infection, it is still plausible that users with higher threat appraisal are more interested in their own risk status and, therefore, more likely to use a DCT app (eg, to be able to detect and react to an infection as early as possible). Therefore, threat appraisal may influence use intention independent of the specified design of DCT apps. In addition, recent research has also shown that the theoretical framework of the HBM does profit from incorporating prosocial aspects of decisions [28,29] (ie, using a DCT app may provide a sense of moral obligation to others). Even though individuals with immunity may perceive a lower personal threat, they may feel a personal obligation to track and inform contacts. Overall, to fully investigate the influence of the perceived usefulness of a DCT system on the use intention, a comparison with system-nonspecific factors (ie, threat appraisal) and personal moral obligation should be made. To the best of our knowledge, no previous study has focused on examining the perception of automation-related usefulness while addressing threat appraisal and moral obligation as system-independent factors influencing use intention.

XSL•FO
RenderX

Research Objective

The objective of this research was to examine how automation-related user experience affects the perceived usefulness of contact tracing as well as use intention of DCT apps and how user experience could be improved. To do so, our approach consisted of multiple methods. The first was quantitatively assessing and analyzing the impact of automation-related user experience (ie, experienced system traceability and perceived trustworthiness) as well as system knowledge on the intention of using a DCT app. The second was contextualizing the effects of automation-related user experience measures with factors related to health protection behavior (ie, threat appraisal and moral obligation). The third was a qualitative analysis of user demands for improved information communication between users and the DCT app. Therefore, the key contribution of this research is a better understanding of how system characteristics lead to perceived usefulness of DCT and how optimal DCT apps can increase use intention through automation-related user experience. Thus, this research supports the human-centered design of DCT apps.

To address these research objectives, 317 users of the CWA DCT system were surveyed about their experience with the app through a web-based questionnaire. A partial least squares structural equation model (PLS-SEM) was used to quantitatively describe the relationships among psychological factors regarding DCT use. This approach was supplemented by a thematic analysis of qualitative user requests on desired communication of information between users and the system.

Related Research

Use Intention of DCT

DCT describes software applications that support documenting information of physical contact or proximity between people (cf [30]). This includes both the (partially) automated acquisition of contact information and the analysis of this information (eg, to determine an individual's risk of infection [31]). In pandemic situations, users might have the goal to avoid contributing to the further spread of the pandemic disease and, thus, face a control task. This means that users need to constantly self-regulate their actions in relation to their environment (eg, how many people around them are infected). While users strive to achieve this goal, they are constantly facing a changing environment (ie, exposure to infected persons). To maintain control, they need to constantly acquire and analyze information and decide, for example, whether they want to isolate themselves. Such actions taken by users have a profound impact on the trajectory of their individual situation—they potentially curtail further contacts and, thereby, change the future information acquisition process. In this process, DCT constitutes a crucial tool for behavioral control as the information provided functions both as feedback for previous behavior and as an indicator for future behavior.

Although DCT applications, especially on mobile devices, first generated high interest during the COVID-19 pandemic [32], they had already been used previously (refer to, eg, the study by Sacks et al [33]). Due to their wide applicability and potential role in public health systems during the COVID-19 pandemic, research on user behavior toward DCT has increased. Here, diverging acceptance models (such as the Unified Theory of Acceptance and Use of Technology and the technology acceptance model) have been evaluated to understand DCT use intention (eg, the study by Velicia-Martin et al [34]).

As indicated at the outset, previous research on DCT app use has leveraged not only acceptance models but also more general models of health behavior such as the HBM or the Theory of Planned Behavior [35]. Such models have been successfully used in research on the uptake and maintenance of other pandemic protective behaviors. In that context, there is consistent evidence of the importance of factors related to the behavior itself, such as perceived usefulness; factors related to perceived risk, such as threat appraisal; and social and normative factors [11,36]. However, in the DCT context, results are mixed. While there is broad support for the importance of factors such as use intention [35] and perceived usefulness [7], evidence of the role of the other factors is less consistent. For example, Tomczyk et al [35] found evidence of the role of both subjective norms and threat appraisal. In contrast, Walrave et al [7] did not include normative factors in their study and found no significant relationship between threat appraisal and DCT adoption. In a different approach to conceptualizing norms, Zabel et al [37] found a strong association between DCT adoption and moral intensity, a construct that derives the perceived obligation for DCT adoption from a range of beliefs, including beliefs about both usefulness and risk. This not only mirrors findings on the association between moral obligation and other pandemic protective behaviors, but as the community benefit of DCT might outweigh the individual benefit, it also appears to be a promising avenue for exploring the relationship between norms and DCT use. Accordingly, it remains an important task of DCT research to understand the relative influence and interplay of both factors such as perceived usefulness, and factors such as threat appraisal or moral obligation on use intention.

One reason for the ambiguity of existing results can be the variability of operationalizations—trust, for example, is highlighted in multiple studies as decisive for DCT use intention [7,35,37]. However, the conceptualization of trust can be challenging and context-dependent [38]. In DCT, for example, trust could influence one's belief regarding how effectively DCT can support the individual in avoiding an infection. On the other hand, trust can be related to the data security of private information (refer to, eg, the study by Altmann et al [39]). Therefore, a context-sensitive and theory-based conceptualization of trust is necessary to operationalize it adequately.

Breaking Down Automation-Related User Experience in DCT

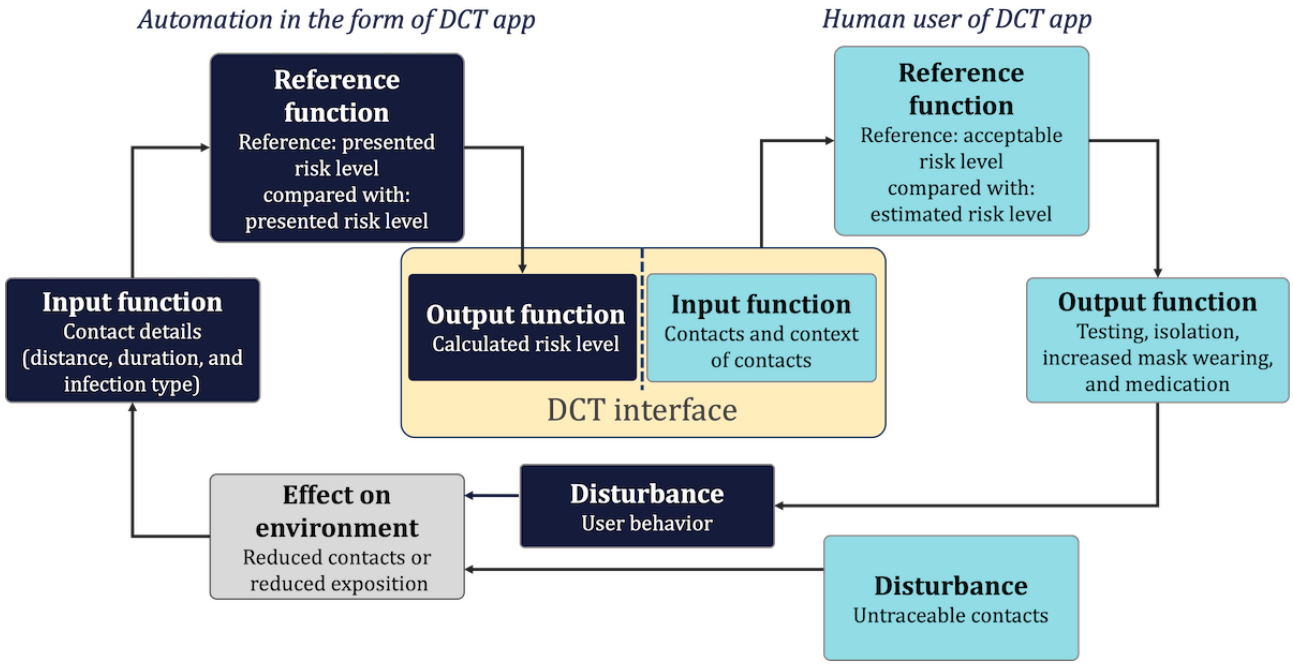
In a pandemic context, the goal of users can be characterized as behavior that avoids both becoming infected and spreading infection to others. Still, they may desire to meet other people or use public transport and, therefore, are continuously adapting their behavior based on how they perceive the risk situation (ie, for simplification, a perceived risk level; refer to the study by Wilde [40]). This risk level refers to the probability of being

XSL•FO
RenderX

infected by, for example, a virus. Acquisition of information on the current risk level is supported by DCT and becomes critical information for comparison, prompting actions to reduce risk.

Contact tracing involves data gathering but also decision-making processes that influence individual and collective health outcomes. It integrates continuous information processing and, therefore, can be viewed through the theoretical lens of control-theoretical conceptions of human-machine systems. The control loop model of action regulation in contact tracing can be extended to accommodate for DCT as automation (ie, a system) that takes over tasks in the acquisition, analysis, and decision selection of contact information [20]. However, maintaining an acceptable risk level [40] is not a singular, finite process but a continuous one. Accordingly, we propose to model information acquisition, analysis, and decision selection as parts of an action regulation consisting of an input function, a reference function, and an output function. As depicted in Figure 1, both human and machine information processing can be modeled within a conceptual control loop to reflect continuous information processing. The conceptual control loop model (Figure 1) illustrates the integration of human and automation activities into a joint action regulation.

Figure 1. Conceptual control loop model of joint human-machine action regulation in digital contact tracing (DCT). The assessment of the machine processing steps (input, reference, and output) is central to the perceived trustworthiness (perceived data validity, perceived goal congruity, and perceived result diagnosticity) of the system.



Based on the model presented in Figure 1, we assumed that users' interaction with DCT apps is based on their evaluation of automated input, reference, and output functions. They assess the correctness of the data that the DCT system uses (*input function*), the data's congruence with the users' goals (*reference function*), and the utility of the data's communicated results (*output function*). Any lack of transparency in their joint action regulation can diminish perceived trustworthiness as well as hamper situation awareness. For instance, if the system fails to capture necessary data accurately or align with personal goals such as identifying the source of infection versus alerting those potentially infected, perceived trustworthiness may decline. Accordingly, parallel to similar phenomena in other automation contexts that do not reveal which information is used as part of the input function, an out-of-the-loop unfamiliarity might cause decreasing situation awareness [20]. Furthermore, the user experience may suffer if the system's output, such as an imprecise infection risk description, is insufficient for users to decide the next course of action, therefore impeding the perceived usefulness.

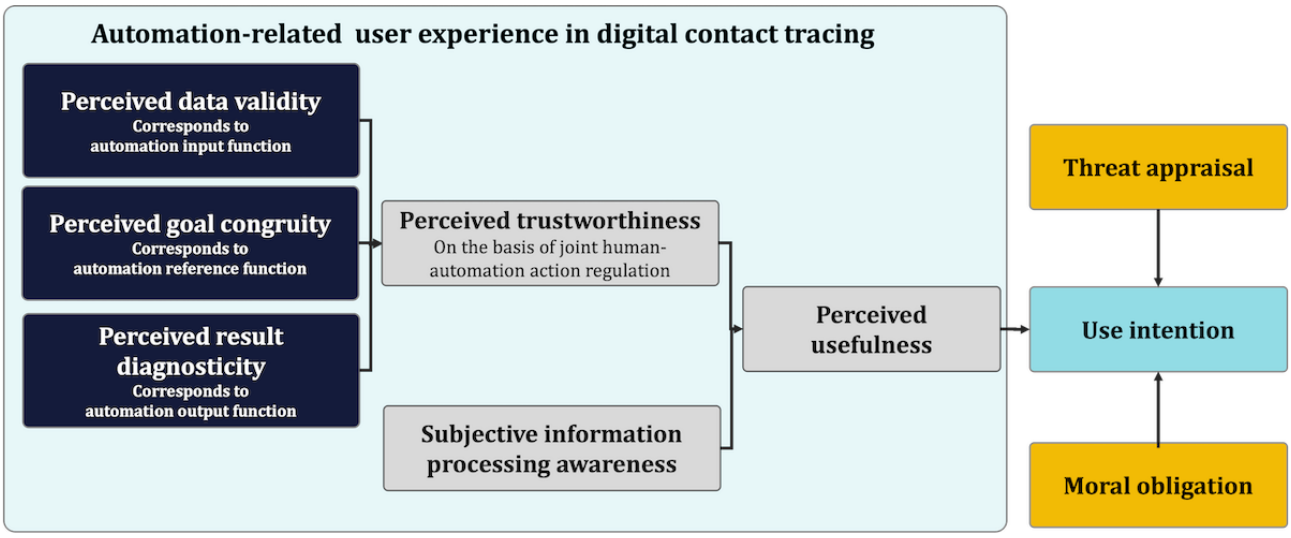
In addition, users' perception of the system is dependent on their expectations of information processing (cf [41]; ie, how the DCT system processes contact-related data). For example, whether a DCT app processes others' vaccination status will only matter to users who are interested in that information, and disclosing that the app processes vaccination information will only impact the system perception of those users. As such, to understand the formation of perceived usefulness, users' subjective situation awareness is more important than their factual situation awareness. However, as introduced by Schrills and Franke [42], subjective evaluation of a user's ability to "perceive, understand and predict a system's information processing," described as subjective information processing awareness, can serve as a construct to assess users' perception of an automation's effect on situation awareness. However, users' perception of their information processing awareness might not be reflected in the accuracy of their knowledge about the system's information processing.

The previous concepts of perceived data validity, goal congruity, result diagnosticity, trustworthiness, subjective information processing awareness, and perceived usefulness can be subsumed as automation-related user experience. Automation-related user experience, following the 9241 standard from the International Organization for Standardization, can be

XSL•FO
RenderX

defined as *the perception and response of a person resulting from using or anticipating the use of automated systems*. On the basis of our proposed conception of automation-related user experience, we conceptualized a model of factors of use intention in DCT centered on perceived usefulness of automation as depicted in Figure 2. In addition, threat appraisal and moral obligation as factors independent of DCT use are integrated as measures to evaluate the influence of automation-related user experience on use intention comparatively. Threat appraisal and moral obligation are not connected with properties of the DCT app; that is, they influence whether a user wants to demonstrate behavior to trace contacts but not how useful a specific app is perceived to be.

Figure 2. Research model on automation-related user experience and the effect on use intention of digital contact-tracing apps.



This Study

On the basis of the presented research model, the objective of this study was to investigate how automation-related user experience affects the perceived usefulness of contact tracing as well as the use intention of DCT and how user experience could be improved. We aimed to contribute to research on DCT adoption and use by examining possible pathways to enhance use intention via user experience. On the basis of the proposed research model, we analyzed the following hypotheses: (1) perceived trustworthiness correlates positively with perceived usefulness (hypothesis 1), (2) subjective information processing awareness correlates positively with perceived usefulness (hypothesis 2), and (3) perceived usefulness correlates positively with use intention (hypothesis 3).

In addition, we examined the relationship among all the aforementioned variables in a structural equation modeling (SEM), where we tested automation-related variables as well as variables not related to the specific DCT system: (1) threat appraisal is positively related to use intention (hypothesis 4) and (2) moral obligation is positively related to use intention (hypothesis 5).

Accordingly, the research model depicted in Figure 2 serves as a basis for an SEM analysis that integrated both automation-related user experience and automation-independent variables (threat appraisal and moral obligation).

We supplemented our quantitative findings with qualitative data on the requirements for improved information processing, providing a deeper insight into users' interactions with the app. This mixed methods approach allowed us to uncover underlying patterns and themes that cannot be identified through quantitative data alone, providing a more comprehensive understanding of the user experience.

Methods

Participants

Participants were recruited via social networks (Twitter [subsequently rebranded X] and Facebook), where an image and a link to the study were shared showing a picture of the CWA and asking for participation (ie, our sample was self-selected). The recruitment strategy specifically targeted individuals who had experience using the CWA. Eligibility for the study required participants to be aged ≥18 years and have at least fluent German skills. The study was conducted on the web, with data collection taking place via a web-based questionnaire between June 1, 2022, and July 31, 2022, using LimeSurvey (LimeSurvey GmbH) [43]. We decided not to inquire further about demographic variables to maintain high levels of privacy due to the context of the study (tracking apps).

A total of 317 participants were included in the study (refer to the Data Exclusion section for further details). As user diversity can have a significant impact on the individual user experience and the perceived trustworthiness, we assessed the affinity for technology interaction (ATI) [44]. ATI describes the individual tendency to actively engage in intensive technology interaction. The ATI was measured using a scale validated in various large samples. Our sample ranged from 1 to 6, with an average value of 4.19 (SD 1.26) which was somewhat higher than the value of 3.5 that Franke et al [44] assumed for the general population based on quota sampling. This corresponds with the self-selection of the sample; we can assume that users who installed the CWA may have, in general, a higher level of ATI than the general population.

XSL•FO
RenderX

Ethical Considerations

This study was registered (under 2022-413) at the Ethics Committee of the University of Lübeck. Before participating in the study, individuals received detailed information about the study and provided written consent to partake. For anonymity, no additional demographic data of the users were queried. No financial remuneration was provided for participation.

Scales and Procedure

Overview

To capture the psychological concepts described previously, multiple scales were developed and presented to participants after they provided informed consent. Except for those for experienced system traceability [42], all items were generated by the researchers based on theoretical considerations and discussed within a team of 3 experts in human-machine interaction.

All items used a 6-point Likert response scale (*completely disagree*=1, *largely disagree*=2, *slightly disagree*=3, *slightly agree*=4, *largely agree*=5, and *completely agree*=6), with the only exception being the semantic differential used for perceived usefulness. For all variables except knowledge, a mean score of all items of the scale was calculated and used for further analysis. All the original items were in German and are presented in this manuscript in English.

Use Intention

Use intention was captured using a 3-item scale focusing on participants' intention and future commitment to use the CWA during the pandemic (Multimedia Appendix 1).

Threat Appraisal

A 4-item scale was used aiming to comprehend the participants' perceived risk and concerns related to a possible infection (Multimedia Appendix 1).

Experienced System Traceability

Experienced system traceability was assessed using the 6-item Subjective Information Processing Awareness scale [42] measuring the perceived transparency, understandability, and predictability of information collection and processing by the system (Multimedia Appendix 1).

Moral Obligation

Moral obligation was evaluated using a 3-item scale capturing the participants' sense of responsibility and ethical obligation toward using the CWA (Multimedia Appendix 1).

Perceived Trustworthiness

Perceived trustworthiness was measured across 3 subscales, each addressing the trustworthiness of input, reference, and output in the cybernetic control loop (Multimedia Appendix 1).

Perceived Usefulness

Perceived usefulness was assessed using a semantic differential scale with labels indicative of the perceived efficiency, precision, safety, complexity, and reliability of the system when cooperating with it (for instructions and labels, refer to Multimedia Appendix 1).

Statistical Analysis

Overview

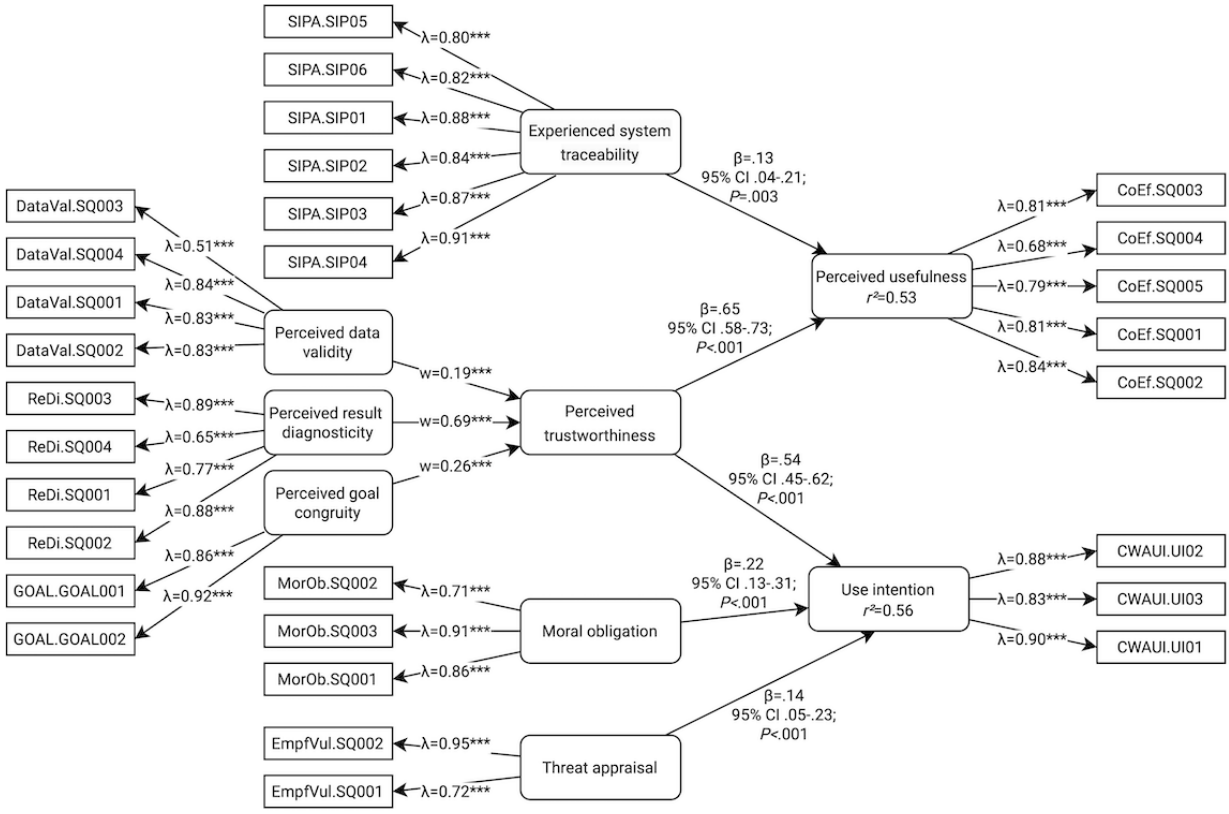
The data collected in this study were analyzed using R (version 4.31; R Foundation for Statistical Computing) [45]. Initially, the normal distribution of the data was tested to ensure that assumptions of normality were met. Given that the data did not follow a normal distribution, nonparametric tests such as the Welch 2-tailed *t* test were applied to determine statistical significance. In addition, considering the multiple comparisons performed in the calculation of correlations, a Bonferroni correction was used to control for the risk of type I error. Corrected *P* values are reported. The analysis was based on the preregistration, which can be found under <omitted for blinded review>.

PLS-SEM is a statistical modeling method combining aspects of regression and factor analysis. It allows for the simultaneous estimation of the relationship between indicators (ie, manifest variables) and constructs (ie, the latent variables formed from the manifest variables) and the relationship between the constructs themselves. These parts of the models are called the measurement model and structural model [46]. PLS-SEM is robust to nonparametric data, can work with small samples, and is especially suited for exploratory research [47], making it a great fit for this study. We followed the extensive iterative process of model assessment described in the work by Hair [46]. Our iterative approach is documented in Multimedia Appendix 2.

The hypothesized PLS-SEM contains all paths depicted in Figure 3. In addition, we tested whether the paths from perceived trustworthiness, system knowledge, and experienced system traceability to use intention were all mediated by perceived usefulness or whether there were also direct effects.

XSL•FO
RenderX

Figure 3. Partial least squares structural equation model after multiple iterations for the proposed research model. Rounded corners indicate constructs based on our research model; rectangular shapes denote indicators that were measured directly in the survey.



All constructs except perceived trustworthiness were specified as mode-A constructs. The respective indicators are described in the Scales and Procedure section. Perceived trustworthiness was specified as a mode-B higher-order construct consisting of perceived data validity, perceived result diagnosticity, and perceived goal congruity. We report explained variance using R^2, path coefficients using β with *P* values and 95% CIs, and effect sizes using the Cohen f^2.

Power

For the PLS-SEM, a retrospective power analysis using the inverse square root method revealed that, given our sample size (N=317), the smallest path coefficient, and a 5% significance level, we achieved a statistical power of 72% [48].

Data Exclusion

Before the statistical analysis, the data set with 370 responses was carefully reviewed for any inconsistencies, missing data, and outliers. Cases with incomplete or implausible responses (53/370, 14.3% in total) were identified and excluded from the analysis to maintain the integrity of the data set.

Qualitative Data Analysis

To obtain a deeper insight into users' demand for information provision and preservation in the interaction with the CWA, qualitative data were collected via open-ended questions (ie, *what information would you like to get from the system?* [Automation to human; question 1] and *What information would you like to feed to the system?* [Human to automation; question 2]).

As a widely used tool, thematic analysis aims to support the systematic identification, analysis, and reporting of patterns (ie, themes) in qualitative reporting data. Both inductive and deductive approaches were applied using theoretical assumptions as the basis for creating the themes, which were then adapted based on the data collected [49]. The data were coded using MAXQDA (version 20; VERBI GmbH [50]). For a structured and reliable analysis approach, a coding scheme with clear definitions of codes and example coding was developed in multiple iterations (Multimedia Appendix 3). For the evaluation, two perspectives of information needs between humans and automation should be covered: (1) human to automation and (2) automation to human. In total, 2 coders coded the data based on the developed scheme. An intercoder reliability of κ=0.90 (for automation-to-human information demands) and κ=0.87 (for human-to-automation information demands) was achieved. Hence, the level of agreement was strong in both cases [51].

Coded themes for information needs in both automation to human and human to automation included contact or risk information, pandemic-related information, app-related information, and assumptions for perceived information processing. Subcodes were created to enhance coding accuracy (Multimedia Appendix 3) but were not analyzed in detail as the focus remained on the top-level codes. Codes that could not be assigned to one of the themes were assigned to the category *others*. As several participants commented, for example, on the suspected reasons for the limitation of information processing, another category was added (ie, assumed reasons for perceived information processing) to avoid losing these data. Both the categories *others* and *assumed reasons for perceived information processing* were not evaluated for this study.

XSL•FO
RenderX

Missing answers to the questions asked and specific statements that there was no demand for information were assigned the code *none*. This code was assigned only once per person and statement. Thus, in the end, it was possible to clearly distinguish how many of the 317 respondents indicated information needs and how many did not. Ultimately, automation-to-human information demand statements from 45.4% (144/317) of the participants and human-to-automation information demand statements from 27.1% (86/317) of the participants were analyzed (Table 1).

Table 1. Number of respondents that indicated information demands versus no information demands.

Variable	Response distribution, n (%)	
	Respondents (n=317)	Responses (n=377)
Demands		
Information demand (A2H[a])	144 (45.4)	257 (68.2)
Information demand (H2A[b])	86 (27.1)	120 (31.8)
No demands		
Information demand (A2H)	173 (54.5)	120 (31.8)
Information demand (H2A)	231 (72.9)	257 (68.2)

[a]A2H: automation to human.

[b]H2A: human to automation.

Results

Overview

For hypothesis 1, the analysis revealed moderate positive correlations for all factors of perceived trustworthiness. The correlation between data validity and perceived usefulness was significant, with a coefficient of r=0.38 and P<.001. The correlation between goal congruity and perceived usefulness showed a coefficient of r=0.47 and P<.001, indicating a moderate positive linear relationship. Result diagnosticity and perceived usefulness exhibited a strong positive correlation, with a coefficient of r=0.56 and P<.001. In general, all measures of perceived trustworthiness and perceived usefulness exhibited a positive relationship, supporting hypothesis 1.

For hypothesis 2, a correlation coefficient of r=0.35 (P<.001) was observed, suggesting a moderate positive linear relationship between subjective information processing awareness and perceived usefulness; a positive relationship between SIPA and perceived usefulness (hypothesis 2) was supported by the data. This indicates that automation-related phenomena such as changes in situation awareness might influence the perceived usefulness of DCT.

For hypothesis 3, the correlation coefficient between perceived usefulness and use intention was r=0.47 and P<.001, indicating a moderate positive correlation. Hence, our results support the hypothesis (hypothesis 3) that perceived usefulness is positively related to use intention (hypothesis 3). In combination with our previous results, this indicates strong relationships between user experience variables and use intention.

In summary, all variables showed statistically significant correlations with perceived usefulness. These correlations ranged from moderate to strong positive relationships. These results strengthen our assumption that perceived usefulness of DCT is strongly related to automation-related user experience.

SEM Approach

The final PLS-SEM is depicted in Figure 3. The explained variance for use intention was R^2=0.56. It was directly predicted by perceived trustworthiness (β=.54, 95% CI .45-.62; P<.001; f^2=0.44), moral obligation (β=.22, 95% CI .13-.31; P<.001; f^2=0.07), and threat appraisal (β=.14, 95% CI .05-.23; P<.001; f^2=0.04). Thus, there was a large effect for perceived trustworthiness and a small effect for the other constructs. Still, hypotheses 4 and 5 were supported.

Within the perceived trustworthiness higher-order construct, the highest weight was assigned to perceived result diagnosticity (w=0.69; P<.001), implying that this subconstruct contributes most to perceived trustworthiness, followed by perceived goal congruity (w=0.26; P<.001) and perceived data validity (w=0.19; P<.001).

We did not find evidence for a mediating effect of perceived usefulness on the paths from perceived trustworthiness, system knowledge, and experienced system traceability to use intention. However, we did find direct effects of perceived trustworthiness (β=.65, 95% CI .58-.73; P<.001; f^2=0.65) and experienced system traceability (β=.13, .04-.21; P=.003; f^2=0.02) on perceived usefulness (R^2=0.53).

Qualitative Analysis

Overview

Two directions of information flow were analyzed to assess the information demands of CWA users: (1) human to automation—information that users want to provide to the system and (2) automation to human—information that users want to receive from the system. In total, 3 overarching themes were explored and analyzed in more detail (Textbox 1).

XSL•FO
RenderX

Textbox 1. Analyzed themes and description of each theme. The detailed coding scheme can be found in Multimedia Appendix 3.

Contact- or risk-related information

- Time-related information: information regarding the period of the contact, the duration of the contact, the time passed since the contact, and the period during which contact tracing was possible
- Location-related information: information related to the place of contact, direct or indirect contact, and indoor or outdoor contact
- Exposition-related information: information about the masking status in the contact situation and the distance between the persons in contact
- Action-related information: information on possible and suggested courses of action after contact
- Information related to the warning person: information concerning the time when the warning person tested positive, the time when the warning person became infected, the warning person's first symptoms, the warning person's vaccination status, and the infected person's virus variant

Pandemic-related information

- Statistics: information related to statistical content on the pandemic in terms of the number of defects or infections

App-related information

- Number of users: information about the number of users of the Corona-Warn-App
- General calculation-related information: information on reasons for changing risk calculation and the system parameters used for calculations
- Certainty about the result: information related to the certainty of the results calculated by the system
- Integration of tests (self- and externally administered): information about the possibility to enter or delete test results on the app
- Linking with private data: information on the possibility of linking app functions with private data

Descriptive Data

Overview

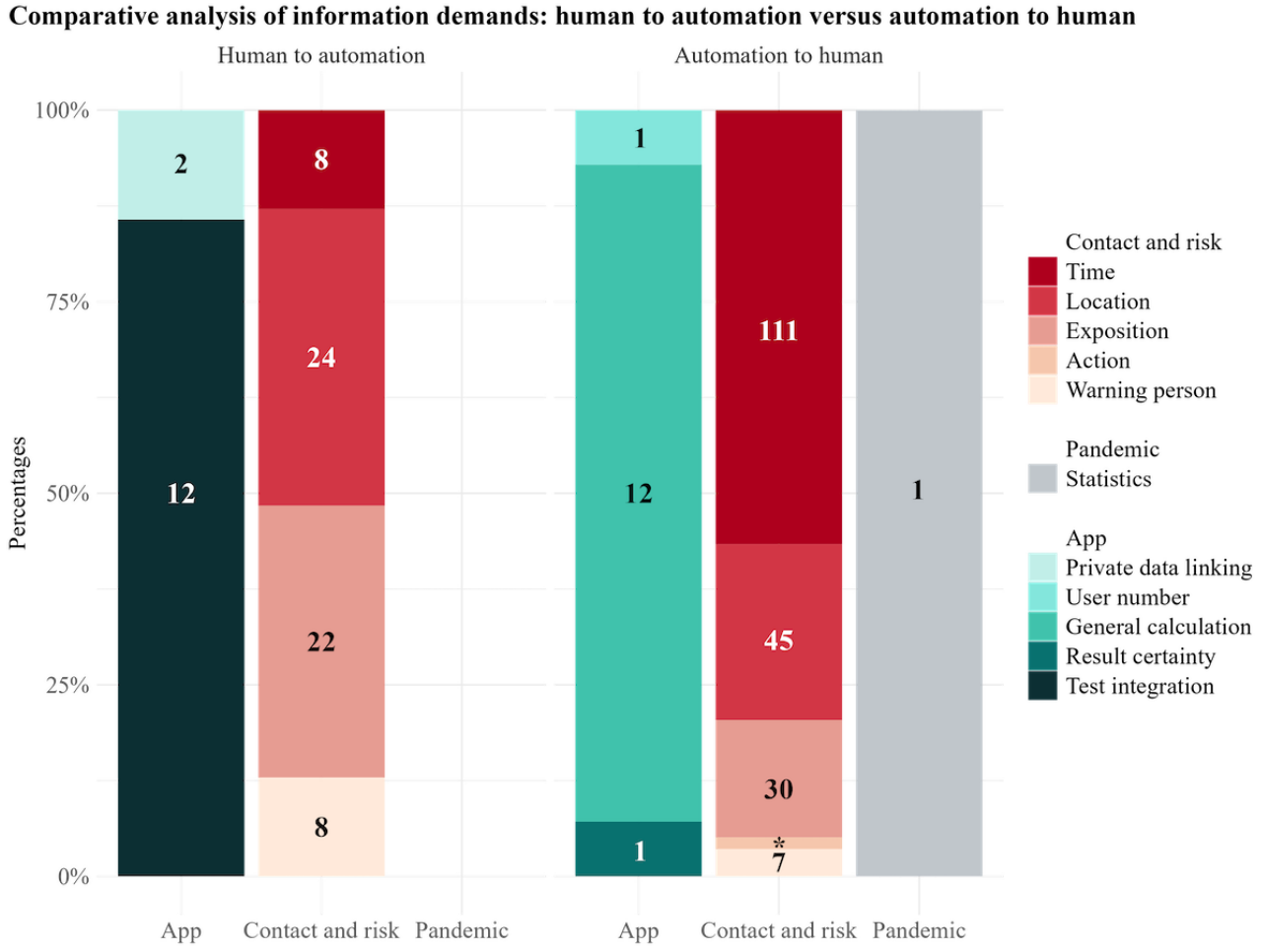
The overall number of statements amounted to 211 in automation to human and 76 in human to automation. Within these 2 categories, the themes were distributed unevenly. Information regarding contact and risk accounted for most statements in both categories (automation to human: 196/211, 92.9% of statements; human to automation: 62/76, 82% of statements). The remaining statements were (almost) exclusively distributed among app-related information (automation to human: 14/211, 6.6% of statements; human to automation: 14/76, 18% of statements) as barely any needs were stated for pandemic-related information (automation to human: 1/211, 0.5% of statements; human to automation: 0 statements).

Regarding the subcodes, the distribution also varied between both themes (Figure 4). For contact- and risk-related information, the information related to time, location, and exposure accounted for the largest proportion of demands within this theme in both categories. However, the distribution of statement proportions differed clearly between automation to human and human to automation. Time-related information was demanded most in automation to human (111/196, 56.6% of statements) but least in human to automation (8/62, 13% of statements). Demands for location-related information did not differ greatly between automation to human (45/196, 23% of statements) and human to automation (24/62, 39% of statements), nor did exposition-related information (automation to human: 30/196, 15.3% of statements; human to automation: 22/62, 35% of statements).

In terms of *app-related information*, the demands for information about the system's general calculation (automation to human: 12/14, 86% of statements; human to automation: 0% of statements) and the integration of tests (automation to human: 0% of statements; human to automation: 12/14, 86% of statements) differed in particular between the categories. The remaining subcodes hardly received any consideration. In both categories (automation to human and human to automation), almost no statements regarding *pandemic-related information* were made.

XSL•FO
RenderX

Figure 4. Relative demands regarding information from automation to human (left) and from human to automation (right). The numbers in the column sections indicate the number of statements under each code.



Human to Automation

In human to automation, certain claims emerged with particular frequency in the demand for contact- and risk-related and app-related information. Information demands on contact and risk mainly focused on time- and exposure-related information. For example, the interest in informing the app of one's location and whether one was in an enclosed space or outdoors was present:

> *Tell the app something about the specific location (enclosed space, fresh air).*

Exposition-related information demands mainly focused on informing the app when one wore or had worn a mask:

> *The wearing of a mouth-nose covering should be entered and thus taken into account in the risk calculation.*

Regarding the demand for the integration of app-related information, the participants predominantly highlighted the integration of self-administered or externally administered tests:

> *That I am Corona positive without having done a Polymera-Chain Reaction (PCR) test. (Perhaps with indication that the result is not PCR verified).*

Automation to Human

In the automation to human category, contact- and risk-related and app-related information were queried with similar frequency. The contact- and risk-related information in this category most often referred to time-related information with a request for the time of the risk encounter. However, the desired preciseness of the temporal data differed (exact time vs more approximate time: "When was the encounter? (At least as a time frame, e.g., between 8-12 o'clock)" vs "The specific time [...] of a risk encounter would be helpful"). The location of the risk encounter was another type of information that participants commonly solicited. Most asked for information about a rather specific location ("At which location did a contact take place?"); few seemed to be interested in the characteristics of the location ("Indoors or outdoors?").

Exposure-related information demanded from the system included the number of devices or persons present at the time of exposure ("[...] with how many devices was the contact?"), the distance to the warning person ("At what distance was the encounter?"), and the masking status. In particular, masking status included the person's own status of having worn a mask or whether the other person was wearing a mask at the time of the risk encounter ("Was I wearing a mask? Was the other person wearing a mask?").

XSL•FO
RenderX

App-related information demands mainly focused on the parameters of the calculation ("What factors led to this result?") and reasons for a status change ("How exactly the risk determination works, i.e., how distance and time to a positively tested person actually have to be, in order for me to receive a notification and for the status to be changed").

Discussion

Principal Findings

The objective of this study was to understand automation-related user experience, its connection to perceived usefulness, and the use intention of DCT. Our data showed that perceived trustworthiness is a critical factor in understanding use intention as well as the perceived usefulness of DCT apps. Interestingly, users' experience of a system as supportive in their action regulation affects their use intention more strongly than external factors such as threat appraisal or moral obligation. In addition, our qualitative analysis revealed that users mainly want to communicate with the system about information that is relevant to their decision-making. For instance, providing more precise information about masking status when in contact with other people could assist a user in making an immediate decision regarding isolation. Overall, our findings suggest a strong relationship between the diagnosticity of automated information processing and use intention.

Practical Implications

As a first major implication, the high effect of result diagnosticity on perceived trustworthiness demonstrates the importance of human-centered information processing in (partially) automated health applications. Within the interconnected human-machine information processing loops (Figure 1), the machine provides information as part of the human input function. As discussed by Miller [52], intelligent systems such as DCT should aim to improve users' ability to access and use (processed) information rather than to present and justify a particular outcome. In DCT app design, the integration of DCT information into a joint human-automation action regulation should be prioritized. Accordingly, when developing evaluative systems [52] that support the evaluation of alternatives rather than suggesting specific actions, it is important to consider what evaluative process a user needs to undertake. While previous research has already identified the need for actionable information [53], the information presented by DCT apps needs to be understood in the context of human action regulation and the influence of automated systems in human action regulation. A possible solution to support diagnosticity in DCT is so-called proactive contact tracing [54], which integrates more information sources and can potentially enrich DCT results.

Second, the results indicate a strong user need for information to be provided in sufficient detail. An interface optimized for communicating information could enable users to make their own assessment of the situation. In many DCT apps, users request the ability to retrieve information about possible contacts, such as time, location, or even the person involved [55]. Our study showed similar results (eg, a high demand for detailed information about the [exact] time of detected contacts). Again, the demand for more detailed information relates to the diagnosticity of the information provided by the system. If users are only given information about their current risk of infection, they cannot evaluate the validity of this information, potentially leading them to ignore it. They would require additional context-related information about potential contacts, such as whether the individuals were wearing masks or were located in an enclosed room, to make informed decisions about their behavior. Our results demonstrate that use intention is strongly connected to the perceived diagnosticity of the DCT app. On the basis of our qualitative findings, we can assume that the diagnosticity of DCT users depends on the level of detail they receive about possible contacts. Accordingly, the provision of details that support users' information processing is even more important for their use intention than threat appraisal or moral obligation. In accordance with psychological research on motivation [56], supporting users' intrinsic motivation for diagnostic information could lead to better adherence regarding DCT apps than, for instance, exposing them to extrinsic motivators that increase threat appraisal (eg, describing the consequences of infection [57]).

Third, in contradiction to users' demand for detailed information on contacts, a major concern in DCT is privacy [55]. While it is often argued that too much detail conflicts with privacy, it is important to find ways to improve the diagnosticity of information as this determines the use intention. Possible solutions include differential privacy, which allows for sufficient detail for increased diagnosticity while keeping personal data confidential. In addition, many users requested features that do not compromise the privacy of others, such as the ability to inform the system about masking status. Thus, allowing users to refine the input received by the DCT app may increase the perceived diagnosticity of the results. The integration of masking status can be seen as a measure to improve the accuracy of the apps in determining risk levels, ultimately increasing the use intention.

Overall, our results suggest that focusing on the diagnosticity of the information presented in DCT apps could result in improvement in users' health behavior. During the COVID-19 pandemic, users reported that they were unsure about the correct or best action to take to contain the pandemic or could not correctly assess the risk of certain situations [58]. However, this certainty is particularly important when it comes to health decisions. With sufficient diagnostic accuracy, DCT apps may be able to better reduce this uncertainty and, thus, become a crucial component in the management of pandemics in the long term, also positively affecting users' willingness to provide data on a social level. It is also crucial that DCT apps do not follow the *recommend and defend* principle [52], which could lead to a long-term reduction in motivation, but instead provide information that supports individual decisions. If compliance with effective pandemic control measures can be increased as a result, it will be possible to respond more effectively to future pandemics.

Theoretical and Methodological Implications

In our data, the perceived trustworthiness of a DCT app had a greater influence on use intention than threat appraisal or moral

XSL•FO
RenderX

obligation. Furthermore, while previous studies [26] have relied on perceived usefulness, our findings in the PLS-SEM do not suggest that it mediates the relationship between perceived trustworthiness and use intention. However, usefulness can be seen as an ambiguous concept without a specific connection to the design of DCT apps. In this way, focusing on perceived usefulness could hinder approaches to improve DCT by adopting DCT app design and functionality. In contrast, a lack of perceived result diagnosticity indicates to developers that the information provided by a DCT app needs to be adapted to have an impact on joint action regulation. Our research suggests that designers of automated systems should specify the potential actions that users can take and identify decision points at which users may require diagnostic information, such as whether to proceed with a specific action. In addition, highlighting the role of diagnosticity indicates how models of technology in medical systems should be developed. Existing models (such as the technology acceptance model) do not specify to what extent a system's usefulness depends on perceived diagnosticity. Our research demonstrates that behavioral models focusing on information-based decisions are needed to address automated technology in health, for example, DCT.

However, one can argue that the difference between perceived result diagnosticity and perceived usefulness is arbitrary; in a joint human-automation action regulation, the diagnosticity of information seems to be equal to perceived usefulness. However, by directly addressing perceived result diagnosticity as a central variable of automation-related user experience, empirical research can identify paths to improve action regulation support of DCT without previously defining what is useful about a system or not. When a DCT app can deliver information that users can use to regulate their actions, users report a higher intention to use it. Therefore, applying result diagnosticity as a variable in human-automation research is a methodological contribution supporting future research in intelligent automation.

On the basis of our findings, future research on DCT needs to determine how to improve the diagnosticity of DCT apps. This paper introduced a conceptual control loop model of joint human-machine action regulation, which can support research approaches in optimizing perceived diagnosticity as a central variable for automation-related user experience. Addressing the joint action regulation in DCT and health behavior is crucial to understand how the information provided by DCT apps can be integrated into human information processing and how DCT apps influence the human output function. Information that improves the evaluation of individual contacts, such as contact location, masking status, or vaccination level, could improve perceived trustworthiness and use intention of DCT apps. By demonstrating how information processing between human users and DCT apps is integrated, our research supports a shift from viewing human users as receivers of machine results to viewing them as actors using DCT information.

All in all, our findings regarding the significance of diagnosticity have implications for the design of automated information processing in a broader context. Users did not primarily prioritize data validity or goal congruence; instead, their focus lay in determining whether they could trust the system to provide information that would assist their own decision-making process. This may be a general trend in automated information processing.

Limitations and Further Research

All participants of this study were users of the CWA. However, as Walrave et al [59] describe, many citizens in Germany did not use DCT apps, for example, because they did not want to share their data or did not think they were effective. Thus, the findings presented on the impact of perceived diagnosticity may not be applicable to citizens who did not use the app at all. These individuals may have chosen not to use the app for reasons beyond those discussed in this paper. The perceived diagnosticity of a DCT app is only relevant for use intention when potential users are interested in determining their individual risk level or making decisions based on their estimated risk level. That is, our sample may bias the results and underestimate factors relevant to nonusers. For example, nonusers might reject the app because they do not trust the provider of the system. Accordingly, the results of our study may support improving DCT for existing users but not convincing nonusers to use DCT. Further studies need to address nonusers and examine how automation-related user experience affects their decision not to use DCT.

In addition, users may have misconceptions about the factors contributing to the risk of infection and may expect the system to provide irrelevant information that does not aid in making an informed decision. Accordingly, they might report a low perceived diagnosticity while the information provided in the app offers sufficient diagnosticity. The accuracy of one's mental model [60] may influence the perception of actual diagnostic information as nondiagnostic (for a discussion of diagnosticity, refer to the study by Garcia-Marques et al [61]). To tackle false models of diagnosticity, DCT apps should support users in correcting their mental model, for example, by explaining how they can use the provided information. This could be done by simulating decision situations with and without DCT information, offering users the experience of diagnosticity.

Improving the perceived diagnosticity could be beneficial for use intention but could negatively affect perceived data privacy [55]. For example, a function that allows users to communicate when they are wearing a mask could be abused to track specific contacts, therefore revealing potential infections of other users. Data privacy is a critical concern in DCT use [59]. Therefore, current DCT apps are designed to protect the data of other users at the cost of the diagnosticity of information. This research did aim to understand the effect of user experience in automated DCT but did not include how users evaluate potential risks of data privacy violations or approaches to address them (cf [62]). Future research should identify how to balance the desired level of perceived result diagnosticity and data privacy concerns. For example, in direct communication, users who reveal information about their web-based status can see the web-based status of others, allowing them to choose which balance between diagnosticity and data protection they desire. The same function could be implemented in DCT apps to support automation-related user experience. Allowing users to choose their level of diagnosticity themselves allows them also to

XSL•FO
RenderX

control how DCT apps influence their decision-making, thus strengthening user autonomy.

Finally, this study had a cross-sectional design that did not assess how automation-related user experience and use intention regarding a DCT app may change over time. Previous research has demonstrated that automation-related user experience can change over time (eg, because users adapt to the system or they improve how they use the system). Future research on automation-related user experience in DCT apps needs to include a longitudinal study design to capture effects of behavior change and users' perception.

Conclusions

In conclusion, this research highlights the relevance of automation-related user experience in DCT and its role in enabling the effective action regulation of DCT users. Here, providing detailed and diagnostic information is crucial for users to make informed assessments of their situation and actions. The presented quantitative results echo the qualitatively assessed user demand for more detailed information about potential contacts, such as time, location, and context (eg, mask use and indoor or outdoor setting).

Interestingly, our data suggest that other factors not directly related to the app, such as moral obligation and threat appraisal, are less relevant compared to automation-related user experience, especially to the perceived diagnosticity of the information provided by DCT apps. The presented results are also more specific than those of previous studies that relied on perceived usefulness. Our research model did not suggest that perceived usefulness mediates the relationship between perceived trustworthiness and use intention. Instead, we propose that DCT designers should focus on providing diagnostic information at critical decision points.

However, privacy remains a major concern in DCT. While it is often argued that too much detail conflicts with privacy, it is crucial to find ways to improve the diagnosticity of information without compromising privacy. Solutions could include differential privacy or features that do not compromise the privacy of others, such as the ability to inform the system about masking status.

The main impact of our results on the design of DCT apps and health policy is that DCT apps need to provide sufficient diagnosticity to be perceived as useful. This means that (1) the possible actions of users need to be understood before the design of the DCT algorithm and apps and (2) the presented information needs to support them in choosing the correct action. Focusing on the diagnosticity of the information presented in DCT apps could, in turn, also influence user performance. During the COVID-19 pandemic, a significant percentage of users reported uncertainty about the best actions to take or could not correctly assess the risk of certain decisions. Therefore, improving diagnostics could contribute to better and safer decisions.

In summary, our study underscores the importance of balancing detailed and diagnostic information with privacy concerns in DCT apps. As we move forward in this digital age, it is crucial to continue exploring ways to optimize DCT while respecting user privacy.

Acknowledgments

This research was funded by the Federal Ministry of Education and Research of Germany within the framework of the Cooperative and Communicating AI project (project 01GP1908).

Authors' Contributions

TS contributed to conceptualization, methodology, investigation, resources, data curation, data analysis, writing—original draft, visualization, and funding acquisition. LK contributed to data curation, data analysis, and writing—review and editing. MG contributed to investigation, data curation, writing—review and editing, and visualization. ACV contributed to writing—review and editing and supervision. TF contributed to conceptualization, writing—review and editing, and supervision.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Scales for the cross-sectional survey study on user experience.
[PDF File (Adobe PDF File), 68 KB - humanfactors_v11i1e53940_app1.pdf]

Multimedia Appendix 2
Documentation of the iterative process of partial least squares structural equation modeling.
[PDF File (Adobe PDF File), 418 KB - humanfactors_v11i1e53940_app2.pdf]

Multimedia Appendix 3
Coding scheme.
[PDF File (Adobe PDF File), 125 KB - humanfactors_v11i1e53940_app3.pdf]

XSL•FO
RenderX

References

1. Barrat A, Cattuto C, Kivelä M, Lehmann S, Saramäki J. Effect of manual and digital contact tracing on COVID-19 outbreaks: a study on empirical contact data. J R Soc Interface 2021 May 05;18(178):20201000 [FREE Full text] [doi: 10.1098/rsif.2020.1000] [Medline: 33947224]
2. Shahroz M, Ahmad F, Younis MS, Ahmad N, Kamel Boulos MN, Vinuesa R, et al. COVID-19 digital contact tracing applications and techniques: a review post initial deployments. Transp Eng 2021 Sep;5:100072. [doi: 10.1016/j.treng.2021.100072]
3. O'Connell J, Abbas M, Beecham S, Buckley J, Chochlov M, Fitzgerald B, et al. Best practice guidance for digital contact tracing apps: a cross-disciplinary review of the literature. JMIR Mhealth Uhealth 2021 Jun 07;9(6):e27753 [FREE Full text] [doi: 10.2196/27753] [Medline: 34003764]
4. Corona-Warn-App. GitHub. 2020. URL: https://github.com/corona-warn-app [accessed 2023-10-16]
5. Corona-Warn-App: 45 millionen downloads. Presse- und Informationsamt der Bundesregierung. 2022. URL: https://www.bundesregierung.de/breg-de/themen/coronavirus/cwa-45-mio-downloads-1994916 [accessed 2023-10-23]
6. Li T, Cobb C, Yang JJ, Baviskar S, Agarwal Y, Li B, et al. What makes people install a COVID-19 contact-tracing app? Understanding the influence of app design and individual difference on contact-tracing app adoption intention. Pervasive Mob Comput 2021 Aug;75:101439 [FREE Full text] [doi: 10.1016/j.pmcj.2021.101439] [Medline: 36569467]
7. Walrave M, Waeterloos C, Ponnet K. Adoption of a contact tracing app for containing COVID-19: a health belief model approach. JMIR Public Health Surveill 2020 Sep 01;6(3):e20572 [FREE Full text] [doi: 10.2196/20572] [Medline: 32755882]
8. Janz NK, Becker MH. The health belief model: a decade later. Health Educ Q 1984 Jan 01;11(1):1-47. [doi: 10.1177/109019818401100101] [Medline: 6392204]
9. Ahadzadeh AS, Pahlevan Sharif S, Ong FS, Khong KW. Integrating health belief model and technology acceptance model: an investigation of health-related internet use. J Med Internet Res 2015 Feb 19;17(2):e45 [FREE Full text] [doi: 10.2196/jmir.3564] [Medline: 25700481]
10. Davis FD. Perceived usefulness, perceived ease of use, and user acceptance of information technology. MIS Q 1989 Sep;13(3):319. [doi: 10.2307/249008]
11. Huang Z, Guo H, Lim HY, Chow A. Determinants of the acceptance and adoption of a digital contact tracing tool during the COVID-19 pandemic in Singapore. Epidemiol Infect 2022 Mar 02;150:e54. [doi: 10.1017/s0950268822000401]
12. Oyibo K, Morita PP. Factors influencing the willingness to download contact tracing apps among the American population. In: Proceedings of the Adjunct Proceedings of the 31st ACM Conference on User Modeling, Adaptation and Personalization. 2023 Presented at: UMAP '23; June 26-29, 2023; Limassol, Cyprus p. 147-156 URL: https://dl.acm.org/doi/abs/10.1145/3563359.3596983 [doi: 10.1145/3563359.3596983]
13. Oyibo K, Sahu KS, Oetomo A, Morita PP. Factors influencing the adoption of contact tracing applications: systematic review and recommendations. Front Digit Health 2022;4:862466 [FREE Full text] [doi: 10.3389/fdgth.2022.862466] [Medline: 35592459]
14. Vogt F, Haire B, Selvey L, Katelaris AL, Kaldor J. Effectiveness evaluation of digital contact tracing for COVID-19 in New South Wales, Australia. Lancet Public Health 2022 Mar;7(3):e250-e258 [FREE Full text] [doi: 10.1016/S2468-2667(22)00010-X] [Medline: 35131045]
15. Prakash AV, Das S, Pillai KR. Understanding digital contact tracing app continuance: insights from India. Health Policy Technol 2021 Dec;10(4):100573. [doi: 10.1016/j.hlpt.2021.100573]
16. Onnasch L, Wickens CD, Li H, Manzey D. Human performance consequences of stages and levels of automation: an integrated meta-analysis. Hum Factors 2014 May 20;56(3):476-488. [doi: 10.1177/0018720813501549] [Medline: 24930170]
17. Bowden V, Griffiths N, Strickland L, Loft S. Detecting a single automation failure: the impact of expected (but not experienced) automation reliability. Hum Factors 2023 Jun;65(4):533-545. [doi: 10.1177/00187208211037188] [Medline: 34375538]
18. Clegg BA, Vieane AZ, Wickens CD, Gutzwiller RS, Sebok AL. The effects of automation-induced complacency on fault diagnosis and management performance in process control. Proc Hum Factors Ergon Soc Annu Meet 2014 Oct 17;58(1):844-848. [doi: 10.1177/1541931214581178]
19. Parasuraman R, Manzey DH. Complacency and bias in human use of automation: an attentional integration. Hum Factors 2010 Jun 20;52(3):381-410. [doi: 10.1177/0018720810376055] [Medline: 21077562]
20. Parasuraman R, Sheridan TB, Wickens CD. A model for types and levels of human interaction with automation. IEEE Trans Syst Man Cybern A Syst Hum 2000 May;30(3):286-297. [doi: 10.1109/3468.844354] [Medline: 11760769]
21. Endsley MR. Automation and situation awareness. In: Parasuraman R, Mouloua M, editors. Automation and Human Performance: Theory and Applications. Boca Raton, FL: CRC Press; 1996:163-181.
22. Bahner JE, Hüper AD, Manzey D. Misuse of automated decision aids: complacency, automation bias and the impact of training experience. Int J Hum Comput Stud 2008 Sep;66(9):688-699. [doi: 10.1016/j.ijhcs.2008.06.001]
23. Volz K, Yang E, Dudley R, Lynch E, Dropps M, Dorneich MC. An evaluation of cognitive skill degradation in information automation. Proc Hum Factors Ergon Soc Annu Meet 2016 Sep 15;60(1):191-195. [doi: 10.1177/1541931213601043]

XSL•FO
RenderX

24. Röttger S, Bali K, Manzey D. Impact of automated decision aids on performance, operator behaviour and workload in a simulated supervisory control task. Ergonomics 2009 May 28;52(5):512-523. [doi: 10.1080/00140130802379129] [Medline: 19296323]
25. Hoff KA, Bashir M. Trust in automation: integrating empirical evidence on factors that influence trust. Hum Factors 2015 May;57(3):407-434. [doi: 10.1177/0018720814547570] [Medline: 25875432]
26. Krüger N, Behne A, Beinke JH, Stibe A, Teuteberg F. Exploring user acceptance determinants of COVID-19-tracing apps to manage the pandemic. Int J Technol Hum Interact 2022;18(1):27. [doi: 10.4018/ijthi.293197]
27. Costa MF. Health belief model for coronavirus infection risk determinants. Rev Saude Publica 2020 May 07;54:47 [FREE Full text] [doi: 10.11606/s1518-8787.2020054002494] [Medline: 32491096]
28. Kojan L, Burbach L, Ziefle M, Calero Valdez A. Perceptions of behaviour efficacy, not perceptions of threat, are drivers of COVID-19 protective behaviour in Germany. Humanit Soc Sci Commun 2022 Mar 24;9(1):97. [doi: 10.1057/S41599-022-01098-4]
29. Chu H, Liu S. Integrating health behavior theories to predict American's intention to receive a COVID-19 vaccine. Patient Educ Couns 2021 Aug;104(8):1878-1886 [FREE Full text] [doi: 10.1016/j.pec.2021.02.031] [Medline: 33632632]
30. Kleinman RA, Merkel C. Digital contact tracing for COVID-19. CMAJ 2020 Jun 15;192(24):E653-E656 [FREE Full text] [doi: 10.1503/cmaj.200922] [Medline: 32461324]
31. Maccari L, Cagno V. Do we need a contact tracing app? Comput Commun 2021 Jan 15;166:9-18 [FREE Full text] [doi: 10.1016/j.comcom.2020.11.007] [Medline: 33235399]
32. Ahmed N, Michelin RA, Xue W, Ruj S, Malaney R, Kanhere SS, et al. A survey of COVID-19 contact tracing apps. IEEE Access 2020;8:134577-134601. [doi: 10.1109/access.2020.3010226]
33. Sacks JA, Zehe E, Redick C, Bah A, Cowger K, Camara M, et al. Introduction of mobile health tools to support Ebola surveillance and contact tracing in Guinea. Glob Health Sci Pract 2015 Nov 12;3(4):646-659 [FREE Full text] [doi: 10.9745/GHSP-D-15-00207] [Medline: 26681710]
34. Velicia-Martin F, Cabrera-Sanchez JP, Gil-Cordero E, Palos-Sanchez PR. Researching COVID-19 tracing app acceptance: incorporating theory from the technological acceptance model. PeerJ Comput Sci 2021;7:e316 [FREE Full text] [doi: 10.7717/peerj-cs.316] [Medline: 33816983]
35. Tomczyk S, Barth S, Schmidt S, Muehlan H. Utilizing health behavior change and technology acceptance models to predict the adoption of COVID-19 contact tracing apps: cross-sectional survey study. J Med Internet Res 2021 May 19;23(5):e25447 [FREE Full text] [doi: 10.2196/25447] [Medline: 33882016]
36. Shushtari ZJ, Salimi Y, Ahmadi S, Rajabi-Gilan N, Shirazikhah M, Biglarian A, et al. Social determinants of adherence to COVID-19 preventive guidelines: a comprehensive review. Osong Public Health Res Perspect 2021 Dec;12(6):346-360 [FREE Full text] [doi: 10.24171/j.phrp.2021.0180] [Medline: 34965686]
37. Zabel S, Schlaile MP, Otto S. Breaking the chain with individual gain? Investigating the moral intensity of COVID-19 digital contact tracing. Comput Human Behav 2023 Jun;143:107699 [FREE Full text] [doi: 10.1016/j.chb.2023.107699] [Medline: 36818428]
38. Johnson M, Bradshaw JM. The role of interdependence in trust. In: Nam CS, Lyons JB, editors. Trust in Human-Robot Interaction. Cambridge, MA: Academic Press; 2021:379-403.
39. Altmann S, Milsom L, Zillessen H, Blasone R, Gerdon F, Bach R, et al. Acceptability of app-based contact tracing for COVID-19: cross-country survey study. JMIR Mhealth Uhealth 2020 Aug 28;8(8):e19857 [FREE Full text] [doi: 10.2196/19857] [Medline: 32759102]
40. Wilde GJ. Risk homeostasis theory: an overview. Inj Prev 1998 Jun 01;4(2):89-91 [FREE Full text] [doi: 10.1136/ip.4.2.89] [Medline: 9666358]
41. Roscoe RD, Wilson J, Johnson AC, Mayra CR. Presentation, expectations, and experience: sources of student perceptions of automated writing evaluation. Comput Human Behav 2017 May;70:207-221. [doi: 10.1016/j.chb.2016.12.076]
42. Schrills T, Franke T. How do users experience traceability of AI systems? Examining subjective information processing awareness in automated insulin delivery (AID) systems. ACM Trans Interact Intell Syst 2023;13(4):1-34 [FREE Full text] [doi: 10.31234/osf.io/3v9b8]
43. LimeSurvey: an open source survey tool. LimeSurvey GmbH. URL: http://www.limesurvey.org [accessed 2024-04-29]
44. Franke T, Attig C, Wessel D. A personal resource for technology interaction: development and validation of the affinity for technology interaction (ATI) scale. Int J Hum Comput Interact 2018 Mar 30;35(6):456-467. [doi: 10.1080/10447318.2018.1456150]
45. R Core Team. R: a language and environment for statistical computing. R Foundation for Statistical Computing. 2021. URL: https://www.R-project.org/ [accessed 2024-04-29]
46. Hair JF, Hult GT, Ringle CM, Sarstedt M. A Primer on Partial Least Squares Structural Equation Modeling (PLS-SEM). 2nd edition. Thousand Oaks, CA: Sage Publications; 2017.
47. Hair JF, Risher JJ, Sarstedt M, Ringle CM. When to use and how to report the results of PLS-SEM. Eur Bus Rev 2019 Jan 14;31(1):2-24. [doi: 10.1108/ebr-11-2018-0203]
48. Kock N, Hadaya P. Minimum sample size estimation in PLS‐SEM: the inverse square root and gamma‐exponential methods. Inf Syst J 2018;28(1):227-261. [doi: 10.1111/isj.12131]

XSL•FO
RenderX

49. Braun V, Clarke V. Using thematic analysis in psychology. Qual Res Psychol 2006 Jan;3(2):77-101. [doi: 10.1191/1478088706qp063oa]
50. All-in-one tool for qualitative data analysis & mixed methods. MAXQDA - Distribution by VERBI GmbH. URL: https://www.maxqda.com [accessed 2022-08-17]
51. McHugh ML. Interrater reliability: the kappa statistic. Biochem Med (Zagreb) 2012;22(3):276-282 [FREE Full text] [Medline: 23092060]
52. Miller T. Explainable AI is dead, long live explainable AI!: hypothesis-driven decision support using evaluative AI. In: Proceedings of the 2023 ACM Conference on Fairness, Accountability, and Transparency. 2023 Presented at: FAccT '23; June 12-15, 2023; Chicago, IL p. 333-342 URL: https://dl.acm.org/doi/10.1145/3593013.3594001 [doi: 10.1145/3593013.3594001]
53. Caballero A, Leath K, Watson J. COVID-19 consumer health information needs improvement to be readable and actionable by high-risk populations. Front Commun 2020 Aug 7;5 [FREE Full text] [doi: 10.3389/fcomm.2020.00056]
54. Gupta P, Maharaj T, Weiss M, Rahaman N, Alsdurf H, Minoyan N, et al. Proactive contact tracing. PLOS Digit Health 2023 Mar 13;2(3):e0000199 [FREE Full text] [doi: 10.1371/journal.pdig.0000199] [Medline: 36913342]
55. Hsu J. The dilemma of contact-tracing apps: can this crucial technology be both effective and private? IEEE Spectr 2020 Oct;57(10):56-59. [doi: 10.1109/mspec.2020.9205550]
56. Martela F, Hankonen N, Ryan RM, Vansteenkiste M. Motivating voluntary compliance to behavioural restrictions: self-determination theory–based checklist of principles for COVID-19 and other emergency communications. Eur J Soc Psychol 2021 Jan 18;32(2):305-347. [doi: 10.1080/10463283.2020.1857082]
57. Morbée S, Vermote B, Waterschoot J, Dieleman L, Soenens B, Van den Bergh O, et al. Adherence to COVID-19 measures: the critical role of autonomous motivation on a short- and long-term basis. Motiv Sci 2021 Dec;7(4):487-496. [doi: 10.1037/mot0000250]
58. Kučera D. Our words in a state of emergency: psychological-linguistic analysis of utterances on the COVID-19 situation in the Czech Republic. Psychol Stud (Mysore) 2021 Jul 13;66(3):239-258 [FREE Full text] [doi: 10.1007/s12646-021-00613-y] [Medline: 34276073]
59. Walrave M, Waeterloos C, Ponnet K. Reasons for nonuse, discontinuation of use, and acceptance of additional functionalities of a COVID-19 contact tracing app: cross-sectional survey study. JMIR Public Health Surveill 2022 Jan 14;8(1):e22113 [FREE Full text] [doi: 10.2196/22113] [Medline: 34794117]
60. Villius Zetterholm M, Lin Y, Jokela P. Digital contact tracing applications during COVID-19: a scoping review about public acceptance. Informatics 2021 Jul 22;8(3):48. [doi: 10.3390/informatics8030048]
61. Garcia-Marques L, Sherman SJ, Palma-Oliveira JM. Hypothesis testing and the perception of diagnosticity. J Exp Soc Psychol 2001 May;37(3):183-200. [doi: 10.1006/jesp.2000.1441]
62. Hang CN, Tsai YZ, Yu PD, Chen J, Tan CW. Privacy-enhancing digital contact tracing with machine learning for pandemic response: a comprehensive review. Big Data Cogn Comput 2023 Jun 01;7(2):108. [doi: 10.3390/bdcc7020108]

Abbreviations

ATI: affinity for technology interaction
CWA: Corona-Warn-App
DCT: digital contact tracing
HBM: Health Belief Model
PLS-SEM: partial least squares structural equation model
SEM: structural equation modeling

Edited by A Kushniruk; submitted 25.10.23; peer-reviewed by H Kang, CN Hang; comments to author 09.01.24; revised version received 12.03.24; accepted 07.04.24; published 25.06.24.

Please cite as:
Schrills T, Kojan L, Gruner M, Calero Valdez A, Franke T
Effects of User Experience in Automated Information Processing on Perceived Usefulness of Digital Contact-Tracing Apps: Cross-Sectional Survey Study
JMIR Hum Factors 2024;11:e53940
URL: https://humanfactors.jmir.org/2024/1/e53940
doi:10.2196/53940
PMID:

XSL•FO
RenderX

©Tim Schrills, Lilian Kojan, Marthe Gruner, André Calero Valdez, Thomas Franke. Originally published in JMIR Human Factors (https://humanfactors.jmir.org), 25.06.2024. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on https://humanfactors.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Health Care Professionals' Perspectives Before and After Use of eDialogue for Team-Based Digital Communication Across Settings: Qualitative Study

Lili Worre Høpfner Jensen[1*], BSN, MSc; Ole Rahbek[1*], MD, PhD; Rikke Emilie Kildahl Lauritsen[1*], BPT, MSc; Søren Kold[1*], MD, PhD; Birthe Dinesen[2*], MSc, PhD

[1]Interdisciplinary Othopaedics, Orthopaedic Surgery Department, Aalborg University Hospital, Aalborg, Denmark

[2]Laboratory for Welfare Technologies-Digital Health and Rehabilitation, ExerciseTech, Department of Health Science and Technology, Aalborg University, Aalborg, Denmark

[*]all authors contributed equally

Corresponding Author:
Lili Worre Høpfner Jensen, BSN, MSc
Interdisciplinary Othopaedics
Orthopaedic Surgery Department
Aalborg University Hospital
Hobrovej 18-22
Aalborg, 9000
Denmark
Phone: 45 60229406
Email: lili.jensen@rn.dk

Abstract

Background: Orthopedic surgical treatment is a transversal task that requires the active involvement of patients, relatives, and health care professionals (HCPs) across various settings. However, after hospital discharge, communication is challenged and undertaken primarily by phone. New digital communication solutions have the potential to create a space for seamless and patient-centered dialogue across discipline and sector boundaries. When evaluating new communication solutions, knowledge about HCPs' needs and perspectives of use must be explored, as it is they who are responsible for implementing changes in practice.

Objective: This study aimed to (1) investigate HCPs' perceptions of current communication pathways (phase 1) and (2) explore their experiences of using a simple messenger-like solution (eDialogue) for team-based digital communication across settings (phase 2).

Methods: We used a triangulation of qualitative data collection techniques, including document analysis, observations, focus groups, and individual interviews of HCPs before (n=28) and after (n=12) their use of eDialogue. Data collection and analysis were inspired by the Consolidated Framework for Implementation Research (CFIR) to specifically understand facilitators and barriers to implementation as perceived by HCPs.

Results: HCPs perceive current communication pathways as insufficient for both patients and themselves. Phone calls are disruptive, and there is a lack of direct communication modalities when communication crosses sector boundaries. HCPs experienced the use of eDialogue as a quick and easy way for timely interdisciplinary interaction with patients and other HCPs across settings; however, concerns were raised about time consumption.

Conclusions: eDialogue can provide needed support for interdisciplinary and cross-sectoral patient-centered communication. However, future studies of this solution should address its impact and the use of resources.

(JMIR Hum Factors 2024;11:e53391) doi:10.2196/53391

KEYWORDS

CFIR; Consolidated Framework for Implementation Research; digital communication; hospital discharge; implementation science; interdisciplinary communication; orthopedic surgery; patient-provider communication; postoperative care; qualitative research; text messaging

XSL•FO
RenderX

Introduction

Treatment of patients undergoing orthopedic surgery is a cross-disciplinary task formed in partnership with the patient. Communication and collaboration between the patient and different professional groups across various settings are key to achieving quality in patient trajectories and clinical outcomes [1-3]. While hospitalization times are decreasing, an increasing part of the postoperative period takes place in the patient's home and with support from municipal health care professionals (HCPs) [4,5]. However, they are largely dependent on contact with hospital staff when problems related to treatment and care arise.

In Denmark, the current means of communication between patients undergoing orthopedic surgery and HCPs across sectors is primarily by phone, but the synchronicity of this is inflexible and time-consuming. Moreover, HCPs across sectors communicate through different electronic systems, but without including patients in the dialogues. New communication strategies must aim to provide seamless communication paths that reach beyond the existing silos of the health care system and include patients as partners [6].

Digital patient platforms are being introduced in Denmark [7,8] as well as internationally [9,10]. Patients can receive digital patient education, see test results, and answer questionnaires used by clinicians to tailor treatment plans. In some cases, patients are given the opportunity to send texts in a secured chat to HCPs at the hospital before and after hospitalization in addition to phone calls. Internationally, secure messaging is reported as the most used feature on patient platforms [9]. Even though questions are not limited to nursing tasks, answering the messages is often delegated to nurses in outpatient clinics or wards at the hospital [7,8]. This leads to duplicate work for the nurses, who will act as intermediaries or gatekeepers for the questions that patients might have, in the same way as secretaries are gatekeepers for patient-initiated phone calls. Moreover, HCPs from the municipality are not involved in these digital encounters. Even though the surgeon at the hospital holds the primary responsibility for the orthopedic treatment [11], there are no direct communication modalities available between the patient, surgeon, and HCPs across sectors in the postoperative period. A team-based approach to the use of digital communication, involving the patient and all HCPs in their care team, may improve postdischarge communication and support patients more optimally after surgery and discharge. Our focus for this study was on communication pathways both involving patient-to-provider communication as well as provider-to-provider communication, as this is interwoven and interdependent in clinical practice.

In an exploratory qualitative study, we tested a simple messenger-like solution for team-based digital communication between patients and HCPs across sectors (eDialogue), and the perspectives of patients and their use of the solution have been reported in another study. However, when testing new communication pathways in health care, it is pivotal to explore the perspectives of all end users to identify their needs, motivations, and barriers to use at an early preimplementation stage [12]. Therefore, this study aimed to (1) investigate HCPs' perceptions of current communication pathways with orthopedic surgery patients and collaborating HCPs across sectors, as well as their expectations for eDialogue (phase 1), and (2) explore their experiences of using eDialogue for team-based communication (phase 2).

Methods

Study Design

We used a triangulation of qualitative data collection techniques to understand contextual factors and what opportunities and challenges exist before (phase 1) and after (phase 2) the use of eDialogue. This included document analysis, observations [13], semistructured focus groups [14], and individual interviews [15]. Reporting this study followed the Consolidated Criteria for Reporting Qualitative Research (COREQ) checklist [16].

Theoretical Framework

Conducting this study, we were inspired by the metatheoretical framework and terminology described by Damschroder et al [17]: the Consolidated Framework for Implementation Research (CFIR). The CFIR is widely used in health services research and specifically adapted to understand facilitators and barriers to implementation, even at an early preimplementation stage [17,18]. CFIR is centered around five key domains related to implementation, including (1) the intervention, (2) the inner setting, (3) the outer setting, (4) the individuals involved in the intervention, and (5) the processes conducted to implement the intervention [17]. To each domain belong underlying constructs, which describe factors that can either motivate or hinder implementation [17]. Selected CFIR domains and constructs guided our data collection by informing the interview guides and the observation protocol in combination with exploratory questions. In an inductive-deductive approach, CFIR domains and constructs were used to structure data analysis and the reporting of our findings, while still being open to emerging themes. By using CFIR, we aimed to promote structured knowledge building for future implementation strategies that may encourage the adoption of eDialogue in clinical practice.

Participants and Setting

The study originated from the orthopedic surgery department at Aalborg University Hospital, which is a tertiary hospital in Denmark. The Danish health care system is mainly financed by general taxes and is therefore provided free of charge to individuals. It operates across 3 administrative and political levels, which are the state (national level), the regions (regional level), and the municipalities (local level). Hospital care is provided by the 5 regions of Denmark, and primary care and social services, such as rehabilitation outside hospitals, home nursing, and physiotherapy, are provided by the 98 municipalities of Denmark. Even though there is cofinancing and close collaboration between the regional and local levels, HCPs are employed in different organizations and use different electronic health records. There are defined care pathways for patients in need of treatment and care across settings that outline the tasks of the HCPs employed at the different levels, just as there is legislation that the HCPs must follow. However, major

XSL•FO
RenderX

challenges exist in communication and collaboration across settings, especially related to patients in transitions of care from hospital to home.

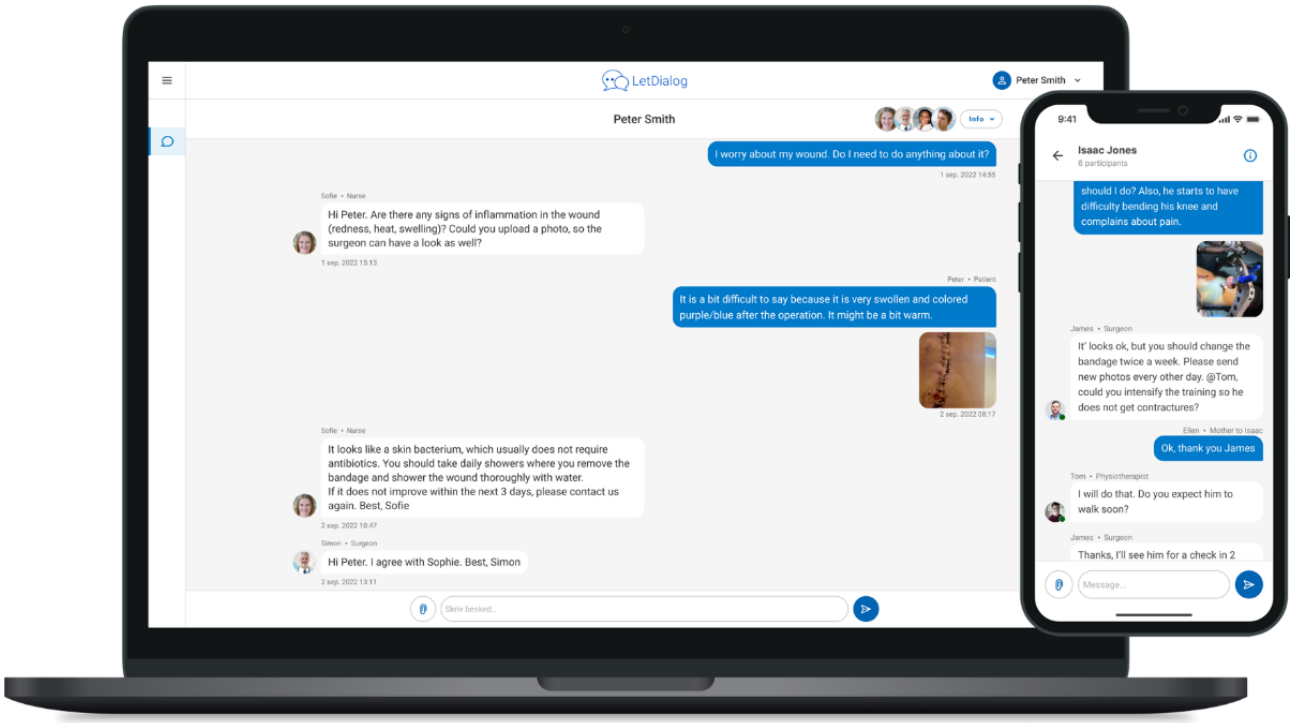
Phase 1: Before eDialogue

In phase 1, orthopedic surgeons, secretaries, nurses, and physiotherapists from Aalborg University Hospital and home care nurses and physiotherapists from the Aalborg municipality were recruited for preintervention focus groups (n=6) to investigate their perceptions of current communication pathways and their expectations for eDialogue. Inclusion criteria were HCPs working with orthopedic patients from 2 different subspecialties that were recruited to test and explore eDialogue. These were patients undergoing either deformity correction surgery involving complex prolonged treatment with hospitalization or anterior cruciate ligament reconstruction performed as day surgery (ie, discharged on the day of surgery). HCPs were recruited from different units at the hospital, including the outpatient clinic, the ward, and the physiotherapy department, and from different districts of the Aalborg municipality. Exclusion criteria were HCPs who had sparse knowledge of orthopedic treatment and care; for example, personnel hired within the past year. We purposely strove to include HCPs from various vocational roles to achieve a detailed understanding of the clinical trajectory and interdisciplinary communication with patients undergoing orthopedic surgery. Inclusion persisted until data saturation was reached for the interviews, that is, no new themes occurred [15].

Intervention: eDialogue

Team-based digital communication between patients and HCPs across settings was facilitated through a technical General Data Protection Regulation-compliant solution assessed by an app for a smartphone or through a website (Figure 1). The technical solution is already in use in some municipalities in Denmark in the field of social education [19], but has never been used to facilitate communication in health care or across sectors. The solution was chosen by the research team before the study based on the simple and intuitive interfaces and discussed with patients undergoing orthopedic surgery and HCPs in an initial workshop before this study.

Figure 1. The figure shows screenshots of digital dialogues between patients and health care professionals (HCPs) across settings from the study. Access was either by app on a smartphone or by web, using a simple messenger-like user interface.



Patients from 2 orthopedic surgery subspecialties were recruited consecutively for this study and offered to use eDialogue for 2 months after they had been discharged with their team of HCPs across settings.

Just as patients were helped to create an account using a digital signature (NemID), HCPs were guided to become users of eDialogue. Most HCPs accessed it through the website, but some preferred access through the app on their smartphones. Finger touch or face recognition could be used for login if access was through the app. During registration, all participants were given a short introduction to how to use eDialogue, including how to send texts and photos and get notifications of new posts. It was explained to HCPs that they were expected to provide answers to patients' questions with a maximum response time of 24 hours on weekdays. In each individual case, patients decided which of the HCPs in their team of care they wanted to join the digital dialogue, and the HCPs were contacted and invited to join by the primary author (LWHJ). All communication was asynchronous, using text messages and photos; thus, no video calls could be made through the solution. Patients had access for 2 months after hospital discharge. Upon request and agreement with their team of HCPs, access could be extended beyond the study period. The digital dialogues were

XSL•FO
RenderX

stored in a secure cloud-based solution [19], and a data processor agreement was made before the study.

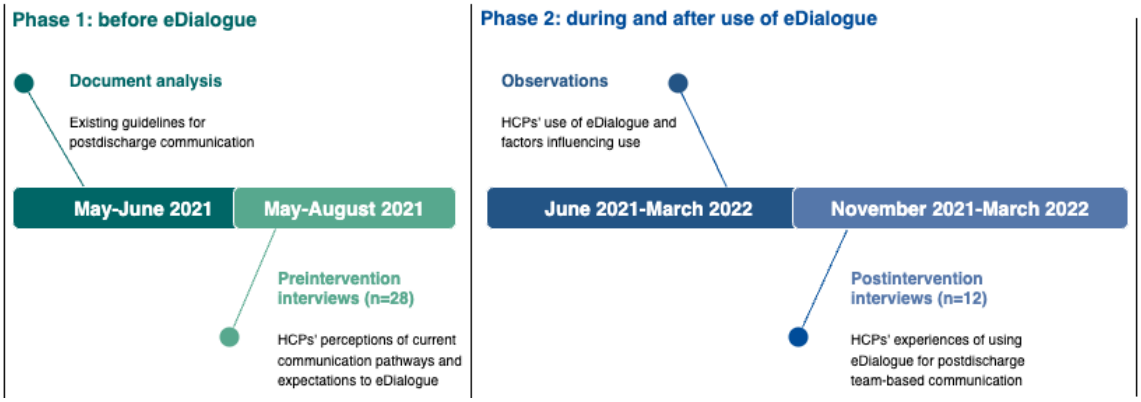
Phase 2: During and After Use of eDialogue

HCPs were recruited for interviews after their use of eDialogue with patients and other HCPs across disciplines and sectors. The inclusion criteria were involvement in eDialogue with ≥3 patients. There were no exclusion criteria.

Data Collection

Data collection was structured according to the two phases to achieve thorough insight into HCPs' perceptions of current communication pathways and their expectations of eDialogue before use (phase 1), and to explore their experiences with access to eDialogue (phase 2). Figure 2 illustrates the triangulation of data collection techniques across the 2 phases of this study.

Figure 2. The figure illustrates the timeline and data collection for phases 1 and 2 of this study. In phase 1, document analysis and preintervention interviews with health care professionals (HCPs) were performed. In phase 2, observations were conducted continuously as eDialogue was used, and, ultimately, postintervention interviews were performed.



Phase 1: Document Analysis and Preintervention Interviews

An initial document analysis of existing guidelines for communication between patients undergoing orthopedic surgery and HCPs across sectors was carried out with the aim of gaining insight into the current context for communication. First, we identified relevant practical documents by searching different Danish web pages related to the political and regulatory guidelines on transitions of care from hospital to home and strategies for using information technologies in health care, for example, the Ministry of Health, the Local Government of Denmark, and the Danish Society for Patient Safety. We also searched the local web page of Aalborg University Hospital for clinical practice guidelines describing the procedures that HCPs must follow when patients or municipal providers contact them regarding discharged patients. Second, we applied a snowball strategy, using references from the initial search. We did not formally analyze the documents, but we used knowledge of the context to understand the framework under which HCPs must work and to qualify the interview guide.

This was followed by focus groups with HCPs across the hospital and municipality (n=28). The aim of the focus groups (n=6) was to explore HCPs' perceptions of current communication pathways and their expectations of eDialogue before use.

The interview guide was inspired by the CFIR Interview Guide Tool [20], including exploratory questions to provide space for emerging reflections. The interview guide was tested on 2 HCPs from the hospital and discussed among the authors until agreement was reached. Minor additions were made before the first focus group.

All preintervention interviews were conducted as semistructured focus groups, dividing HCPs according to their vocational roles and setting (hospital or municipality). HCPs interviewed were surgeons at the hospital (n=5), secretaries from the hospital (n=3), nurses from the hospital ward (n=5), nurses from the outpatient clinic (n=3), home care nurses from the municipality (n=3), physiotherapists from the hospital (n=5), and physiotherapists from the municipality (n=4). Using preexisting groups as focus groups was based on the assumption that it would make participants discuss and compare their reflections in depth in the same context and without an underlying power structure that could occur if professions were mixed [14]. At the beginning of each interview, background variables such as gender, vocational role, and years of experience with patients undergoing orthopedic surgery were collected.

All interviews with HCPs from the hospital were conducted face-to-face by the first author (LWHJ). For the first 2 focus groups, a project nurse was present to register observations during the interviews and to take notes to qualify and supplement the interview. Focus groups with HCPs from the municipality were performed remotely by video, as data collection occurred during the coronavirus outbreak and most HCPs outside of the hospital were not physically located in the same place. The interviewer summarized key points during and at the end of each focus group to facilitate further reflection and to make sure her interpretation corresponded with what the HCPs had said [15]. Field notes were made at the end of each focus group so as to remember details of the context, group interaction, and nonverbal communication [15]. The focus groups lasted an average of 1 hour (between 45 and 90 minutes).

Phase 2: Observations and Postintervention Interviews

In total, eDialogue was used with 31 patients and with the involvement of 24 different HCPs. When the last patient had

XSL•FO
RenderX

had access to eDialogue with their team of HCPs for 2 months, a convenience sample of participating HCPs across the hospital and municipality were interviewed (n=12), including surgeons from the hospital (n=5), physiotherapists from the hospital (n=2), and from the municipality (n=5). We performed 7 individual interviews with physiotherapists across hospitals and municipalities and 1 focus group with 5 surgeons. The aim of the interviews was to explore their experiences with eDialogue. All interviews were conducted by LWHJ, audio recorded, and followed a predefined semistructured interview guide inspired by the CFIR Interview Guide Tool [20] and additional exploratory questions. Interviews with HCPs from the hospital were performed face-to-face, and interviews with HCPs from the municipality were conducted remotely based on the participants' wishes.

During the study period, we observed the use of eDialogue by HCPs and documented this in Word (Microsoft Corporation) files. The aim was to observe issues related to HCPs' use of eDialogue that were reported to the project group or observed in dialogues (an administrator from the project group was present in all dialogues to observe if eDialogue was used in acute situations). HCPs were encouraged to contact the first author if they experienced any problems with eDialogue or had concerns or questions during use, and these were documented as well. Data collected through observations were used to qualify the follow-up interviews in phase 2 and were also imported to NVivo (QSR International) for analysis in conjunction with interview data.

Data Analysis

Data were analyzed for phase 1 and then phase 2, respectively. Interviews were audio recorded using a digital voice recorder (DM-450; Olympus) and transcribed verbatim immediately afterward. Word files with the transcriptions were imported to NVivo for data analysis (NVivo 12, version 20.6.2) [21]. Inspired by Brinkmann and Kvale [15], using an inductive-deductive approach, we performed thematic analysis focusing on meaning (Textbox 1).

Textbox 1. Steps of the thematic analysis using an inductive-deductive approach.

Meaning coding

- Full transcripts were read several times by both LWHJ and REKL.
- To define the initial coding template and to achieve intersubjectivity, the first 4 interviews of each phase were coded by LWHJ and REKL individually before meeting to compare and discuss codes until mutual agreement was achieved. When the coding template was defined, LWHJ applied the same codes to the entire data set. The approach to this step was inductive, thus reflective of the issues raised in the data set.

Meaning condensation

- Theme development was undertaken with a more deductive approach, where domains and constructs from the Consolidated Framework for Implementation Research (CFIR) were used to organize the codes and inform theme development to specifically focus on facilitators and barriers to eDialogue. However, in developing themes, we were open to emerging themes that did not fit the CFIR domains and constructs. Codes and themes were reread and revised by LWHJ in collaboration with REKL and BD in several iterations.

Meaning interpretation

- Definitions and narrative descriptions of themes were made. Data extracts were selected to be presented in the manuscript.
- The final analysis and description of the findings were written.

Data analysis was conducted separately for phase 1 and phase 2 following the 3 steps of meaning coding, meaning condensation, and meaning interpretation. In phase 2, we added notes from observations to the data set to achieve an in-depth understanding of the context in which HCPs had used eDialogue and any problems occurring during use.

Ethical Considerations

The Ethics Committee of Northern Jutland was contacted before the start of the study. They decided by email on March 18, 2021, that the study did not require approval (journal number 2021-000438), as the intervention would not have consequences for diagnostics or treatment. We registered the study at the Regional Committee on Health Research (ID 2021-057). The study followed the Helsinki Declaration, and all participants received both oral and written information as well as thorough guidance in the use of eDialogue. To take into account patients' possible use of eDialogue in emergency situations, an administrator was present in all digital dialogues.

Results

Participant Characteristics

In phase 1, a total of 28 HCPs were recruited across vocational roles and hospital and municipal settings (Table 1). All surgeons, nurses, physiotherapists, and secretaries from the clinical orthopedic surgery subspecialties at the hospital, from which the patients were recruited (deformity correction or anterior cruciate ligament injury), were invited to participate in interviews. However, 2 surgeons, 1 nurse from the outpatient clinic, 1 nurse from the municipality, and 3 secretaries were not able to. Nurses from the ward were purposefully selected based on years of experience and a pragmatic approach to who would be able to participate in interviews during their work hours. On average, HCPs had 11 (range 1-30) years of experience with patients undergoing orthopedic surgery.

XSL•FO
RenderX

Table 1. Vocational roles of health care professionals who were interviewed in phases 1 and 2.

Vocational role	Phase 1 (N=28), n	Involved in eDialogue (n=24), n	Phase 2 (n=12), n
Orthopedic surgeon, hospital	5	7	5
Nurse, outpatient clinic, hospital	3	1	N/A[a]
Nurse, ward, hospital	5	N/A	N/A
Physiotherapist, hospital	5	5	2
Secretary, hospital	3	N/A	N/A
Physiotherapist, municipality	4	11	5
Nurse, municipality	3	N/A	N/A

[a]N/A: not applicable.

In phase 2, a total of 12 HCPs were included for interviews, of whom 8 had also participated in focus groups in phase 1. The HCPs recruited at this stage were a sample of those who had experiences with communication in eDialogue (Table 1). Of whom, 8 HCPs interviewed for phase 2 had also participated in focus groups in phase 1. In total, 24 HCPs across the hospital and municipality were involved in eDialogue. However, we prioritized including those who had been set up to communicate in eDialogue with ≥3 patients. One nurse from the outpatient clinic had been involved in 3 dialogues but was not able to participate due to being absent at the time of the interviews. No nurses from the ward or the municipality were users of eDialogue and thus were not interviewed in phase 2. Secretaries were not interviewed in phase 2, as we decided not to include them in eDialogue at this point.

On average, there were 3.3 (range 2-4) HCPs per patient in the dialogues. All patients were at least connected with the orthopedic surgeon, and 25 of 31 patients had their municipal or hospital-based physiotherapist involved as well.

Themes and Subthemes Identified in Phases 1 and 2

In Table 2, the findings of the analysis of phases 1 and 2 are presented together in main themes organized by the CFIR domains and constructs and additional subthemes. This is to display the before-and-after perspectives of HCPs. Following the table, we elaborate on subthemes in narrative text according to phases 1 and 2 and by using selected quotes from interviews.

The main themes are organized by CFIR domains and constructs, and subthemes elaborate on these for phases 1 and 2, respectively. Emerging themes occurred in both phases that did not match any of the CFIR constructs, and they are therefore described under additional emerging themes.

XSL•FO
RenderX

Table 2. Themes and subthemes from phases 1 and 2 organized by the Consolidated Framework for Implementation Research (CFIR) and additional emerging themes.

Main themes: CFIR domains and constructs	Subthemes	
	Phase 1: before eDialogue	Phase 2: after eDialogue
Intervention characteristics		
Relative advantage	• Contradictory expectations for using eDialogue versus phone call • A lifeline and reassurance for both patients and HCPs[a] • Hidden work can become visible	• Quick and easy to interact in eDialogue • Photos in eDialogue improve the quality of communication
Adaptability	• N/A[b]	• Development of individual strategies and workflows for the use of eDialogue
Outer setting		
Needs and resources	• Patients are messengers of information between HCPs	• Timely and effective interdisciplinary communication with patients across settings
Inner setting		
Tension for change	• Feeling like an insufficient intermediary • Phone calls are disruptive, yet necessary	• N/A
Relative priority	• Experiences of technology fatigue	• N/A
Compatibility	• N/A	• Divergent perceptions of how well eDialogue meets needs
Available resources	• N/A	• Concerns about resource consumption • Need for clarification regarding financial incentives
Characteristics of individuals		
Self-efficacy	• N/A	• To express oneself in writing
Additional emerging themes		
Previous experiences with digital communication	• Email and SMS text messaging are already used with patients and for interdisciplinary communication; however, standardization is lacking	• N/A
Reflection and learning	• N/A	• A new space for studying patients' needs

[a]HCP: health care professional.

[b]N/A: not applicable.

Phase 1: Current Communication Pathways and Expectations for eDialogue

Intervention Characteristics

Even though the majority of HCPs expected eDialogue to provide optimized interdisciplinary communication, prevent conflicting recommendations to patients, and provide easier access for patients, there were contradictory expectations for the use of eDialogue. On one hand, HCPs had concerns about whether answering messages would require more of their time and go beyond working hours, but on the other hand, they thought it would be easier to answer in eDialogue than by phone. Concerns also centered around whether using text as a communication medium would be adequate for all patients and if misunderstandings would occur due to wrong interpretations. HCPs were especially worried about whether they would pick up on complications to the same extent as they do by phone.

> *I can't hear the patient's voice answering back, and if they have understood my answers (…) however, it depends on the complexity of their questions, whether it's just how many repetitions was it, or something that could be more serious.* [Physiotherapist, municipality, preintervention interview]

Being able to send photos in eDialogue was expected to be an important feature that might offset the challenges of using text for communication. Even though some HCPs had reservations and conflicting opinions about eDialogue before its use, they all agreed that it would be a reassurance and a lifeline for both patients and HCPs across settings. Additionally, a

XSL•FO
RenderX

physiotherapist from the hospital reflected on how eDialogue might bring "hidden tasks" to light.

> *What I thought at first would be negative like, oh then we have to do that too, will probably actually be reversed, so that the hidden tasks, we solve by calling and writing notes and emails and things like that, becomes more visible and can be accounted for during work hours.* [Physiotherapist, hospital, preintervention interview]

Outer Setting

The analysis revealed that HCPs experience current communication pathways in the postoperative period to be challenging, both by phone and existing electronic systems. This leads to workarounds, such as HCPs giving patients oral instructions or written notes to bring to other HCPs to ensure timely information. However, using patients as messengers of information between HCPs is perceived as insufficient yet necessary in current communication pathways. A physiotherapist from the municipality described how current systems do not support the patient's trajectory across sectors.

> *There are watertight shutters between the communication systems, i.e., what they write in the medical record at the hospital and what I write here. The surgeon at the hospital can't see that, and (…) I can't see his note.* [Physiotherapist, municipality, preintervention interview]

Inner Setting

Across professional roles and settings, HCPs expressed a need for change to enable easier sharing of knowledge and communication. This was especially the case for complex and long-term orthopedic surgical treatments where multiple HCPs are involved. Knowing each other across settings, for example, by being former colleagues, was a mediating factor for communication between HCPs. However, it was not perceived as sustainable.

Getting in contact with each other and patients by phone is considered time consuming due to the synchronicity of phone calls. A nurse from the inpatient department described how phone calls would sometimes be left until the next day if questions required the involvement of another HCP. This left the nurses feeling like inadequate intermediaries and could be a risk to patient safety. Similar experiences were described by physiotherapists, who often found themselves being asked about issues outside of their competencies; for example, questions about wounds and medication.

HCPs from the hospital described how phone calls are disruptive to their work processes, even though they understood the need for them. In addition to inquiries from patients, they receive phone calls from a wide range of HCPs in hospitals, municipalities, and private settings. Although secretaries act as gatekeepers, nurses from the outpatient clinic, and in the inpatient department in particular, handle many phone calls daily.

> *It's constant, isn't it? (...) it takes my attention away from the dialogue, the communication and the relationship that I'm in the middle of. Then you're like, oh sorry, this phone call is actually more important than you are (the patient they are with).* [Nurse, outpatient clinic, hospital, preintervention interview]

Addressing eDialogue as a novel communication solution to support team-based communication between patients and HCPs across settings, most HCPs were positive about the change it might bring. However, they expressed some degree of technology fatigue that made them skeptical of yet another system without integration into existing systems.

Emerging Theme

HCPs described previous experiences with using digital communication with patients, usually by email or SMS text messaging. Most often, it is used as a way to provide psychological reassurance to patients or to solve specific complex problems, where the HCPs have specialist knowledge. Even FaceTime was described as being used once with a patient to inspect a wound from a distance. However, the disadvantages of the current nonsystematic use of digital communication with patients were reflected. Concerns were raised regarding using a private phone number and the risk of introducing data security breaches. Also, giving some patients the opportunity for direct digital contact and others not was perceived as problematic. Thus, if used inconsistently, it may lead to inequality in patients' access to health care.

Furthermore, HCPs described how they use email or SMS text messaging to communicate with each other, for example, to share thoughts on treatments or rehabilitation. They do this as a workaround to traditional communication pathways or because it is perceived as less disturbing to each other. Thus, the use of digital text-based communication is not uncommon for HCPs in this study. However, it is not standardized or even articulated among colleagues or management.

Phase 2: HCPs' Perspectives of eDialogue After Use

Intervention Characteristics

All HCPs agreed that the technical solution for eDialogue was very intuitive and did not need a thorough introduction, as opposed to other solutions with more features. Most HCPs articulated that questions were quick and easy to handle during work hours. Especially the asynchrony of the contact and the use of photos improved the quality of communication and their experiences of eDialogue for patient communication.

> *The big advantage of this, is that they can send a photo (…). If it wasn't a possibility, I think there would be a lot of writing about something that we couldn't really clarify, and then we would still have to call them in (for an extra check). Being able to send a photo, that's really crucial for this to work.* [Surgeon, hospital, postintervention interview]

The analysis demonstrated that HCPs developed individual strategies for answering questions in eDialogue. Notifications were automatically sent to participants when there were new messages in the system, but there were no integrated reminders

XSL•FO
RenderX

to follow up if the messages were not read within 24 hours, and this led to the development of individual workflows.

> *(The notification) on email, when there is a new message, I will not delete it until I have answered. That way, it helps me keep track.* [Surgeon, hospital, postintervention interview]

eDialogue was mainly used by patients as a place to ask postdischarge questions to HCPs. In general, most questions from patients were answered by surgeons and physiotherapists from the hospital. Municipal physiotherapists described being hesitant to involve themselves actively in answering, as they experienced hospital staff being quick to answer the patients. However, they emphasized that they used the information given to the patient by hospital staff in their subsequent contact with patients. This "indirect" use was perceived as valuable to them.

> *It has been very rewarding to just follow the dialogue, even though I was not active in it. The fact that the patient can just send a photo and ask 'what does this look like?', then he is immediately calmed down. It's rather smart, and also that I know of it right away.* [Physiotherapist, municipality, postintervention interview]

Outer Setting

HCPs stressed that the team-based approach made it easier to share timely information with the patient and other HCPs, and thereby it created more effective communication pathways. Physiotherapists highlighted how their previous perceptions of being an insufficient intermediary between the patient and other HCPs were changed when communication could take place directly in eDialogue.

> *It was actually really nice that he (the patient) just took it directly with the surgeon. Because I can have doubts (…) and you don't want to burden the surgeon by calling.* [Physiotherapist, municipality, postintervention interviews]

Inner Setting

Even though HCPs acknowledged the impact that eDialogue had for patients, there were discrepancies in their perception of how it was used in this study, and it affected their acceptance of the solution. For example, some HCPs thought that the team-based approach was not necessary for all patients involved or that they lacked a secretary for administrative tasks. As such, they highlighted that some questions might be better answered in other ways, for example, by providing better patient education or by including other HCPs in the dialogue.

> *I think it is difficult to say that the patients' questions are not relevant because they must be since they ask them, but who should answer them, and how quickly should they have an answer, can be discussed.* [Physiotherapist, hospital, postintervention interview]

However, when using eDialogue with patients for complex orthopedic treatments, HCPs expressed that the team-based approach was very valuable to the patients and their workflows.

> *I think it was good. They (patients) feel that there is a team around them, and I get the feeling that I'm not the only one being responsible. Also, I don't have to spend time calling the physiotherapist to say 'Hey, can't you just look at this?' when he's already in the dialogue.* [Surgeon, hospital, postintervention interview]

HCPs strongly experienced that access to eDialogue provided reassurance for patients. However, in consideration of the sparse health care resources, it was a general opinion that eDialogue should only be offered to patients for complex treatments. This provoked an ethical discussion of how HCPs could distinguish between who should be offered the solution and who should not. HCPs highlighted that an assessment of effects should be addressed, both in terms of resource consumption and patient outcomes.

> *One of my concerns with systems like this is that if we have to use it with all patients (…), then I think it could become a burden. And also, I think it will be difficult to say, well, it's only for some patients, because why them?* [Surgeon, hospital, postintervention interview]

HCPs agreed that clarification is needed regarding financial incentives before implementing eDialogue. Along with concerns about resources to answer the questions, this was a perceived barrier to use.

> *I think the barriers are time and finances (...) there is, of course, someone who looks at what I produce. And I think it should be some kind of service that should be visible (to others), if we have to evaluate a photo or send back a response (through eDialogue).* [Physiotherapist, municipality, postintervention interview]

Characteristics of Individuals

In all interviews, HCPs had concerns about whether they expressed themselves clearly enough in writing and how their "tone of voice" would be perceived by patients when formulated in texts. In reflection, they emphasized that the same concerns could arise when talking to patients on the phone.

> *Regardless of whether it's something you say to them or something you text them, it's just as important that you use words they can understand, and I actually often think it's a little easier when you text because you have time to think about it.* [Surgeon, hospital, postintervention interview]

There were clear differences in how HCPs expressed themselves in the texts, and this was discussed in one of the focus groups, where a surgeon had been involved in another surgeon's dialogue due to vacation.

> *I think he (the other surgeon) is very kind in his feedback. I actually noticed that, you (addressed to the other surgeon) have formulated yourself in such a very friendly way, in contrast to what I did to start with. I made it very short, like I might normally answer a text message with a friend (…). I had to remind myself that they don't know me (…) it might be important to pay attention to that.* [Surgeon, hospital, postintervention interview]

XSL•FO
RenderX

Emerging Theme

Both surgeons and physiotherapists described that using eDialogue created interdisciplinary reflection and learning about patients' needs after discharge, and that frequently asked questions could be used to improve future patient education.

> *It gives feedback in relation to the material we use and the way we inform patients now. It might actually be very nice for all of us to know this.* [Physiotherapist, hospital, postintervention interview]

Ultimately, HCPs pointed out that they could learn from each other by reading each other's answers to patients.

Discussion

Principal Findings and Comparison With Previous Work

This study first investigated HCPs' perceptions of current communication pathways with patients and other HCPs involved in the patient's trajectory after orthopedic surgery and discharge, along with their expectations for eDialogue before its use (phase 1). Following initial document analysis, we included a wide range of HCPs across vocational roles and settings in focus groups to obtain an in-depth understanding of their needs and attitudes toward eDialogue. These perspectives are important to capture, as individual and contextual factors as well as initial perceptions of eDialogue may motivate or hinder use [17]. The findings of phase 1 showed that, on the one hand, HCPs perceived a significant tension for change. Current communication pathways are perceived as insufficient, phone calls are disruptive, and patients unfortunately become messengers of information between HCPs across settings. On the other hand, HCPs expressed conflicting attitudes toward eDialogue in advance of its use. Positive or negative attitudes were not limited to certain vocational roles but were expressed in all groups and also as an internal dilemma inherent to the individual. However, there were clear expectations for eDialogue to support patients in the postoperative period and consensus that it may provide optimized interdisciplinary and cross-sectoral communication. At the same time, HCPs experienced some degree of technology fatigue and significant worry that eDialogue would be time-consuming for them to handle.

Second, we explored HCPs' experiences of using eDialogue for team-based digital communication through observations and postintervention interviews (phase 2). Knowing that, even with highly developed plans for execution, undiscovered factors can undermine implementation efforts in the real world [17,18], we searched to identify facilitators and barriers to implementation from the perspectives of key users at an early stage. Findings from phase 2 showed that HCPs experienced eDialogue as a quick and easy way to interact with patients and other HCPs and that eDialogue could support timely and effective interdisciplinary communication across settings. As such, the positive perceptions of the importance of eDialogue described in the preintervention interviews were maintained. Similarly, the use of photos was expected to be important in preintervention interviews, and in postintervention interviews, photos were even suggested as being a significant quality-enhancing element compared to traditional phone calls. Similar findings have been described in other studies of digital communication in health care [22-24].

In interviews in phase 1, HCPs described that they had concerns about communicating with patients in texts because they feared overlooking an important complication or that the patient would misunderstand their written responses. In phase 2, HCPs still expressed concerns about whether they expressed themselves clearly enough. However, they pointed out that the same risks can be present in phone consultations. This perspective is supported by a recent study of telephone consultations in Denmark. Jensen et al [25] found that communication in consultations concerning back pain preceding out-of-hospital cardiac arrest was influenced by the communicative preconditions of the call-taker, thereby addressing the fact that a meaning-constitution is undertaken in the interaction between the patient and the call-taker, not always reflecting the actual problem. To learn from this, HCPs involved with patients through eDialogue and other digital communication solutions must be aware that communicative interaction is always an interpretative task for the receiver of a message. Even though the HCPs' concern might decrease as they gain more experience communicating in writing, their self-efficacy should be supported by formulating clear recommendations, training, and supervision.

Across the interviews of phases 1 and 2, HCPs expressed concerns regarding resource consumption; this was particularly evident among hospital staff. While acknowledging patients' need for easier access to communication with HCPs after discharge, HCPs questioned if the team-based approach was necessary for patients undergoing less complex orthopedic treatments. Nevertheless, there was consensus that eDialogue can support patients in complex and long-term treatments and that a needs assessment to learn who will benefit the most from eDialogue should be made before its implementation so as to best match resources with actual needs. Other studies investigating the use of team-based digital communication have primarily focused on patients with cancer or chronic diseases [26-29]. Patients undergoing orthopedic surgery for complex and long-term treatment suffer similar challenges in health care communication [30], and therefore it is also relevant to develop and test solutions for this group. By using eDialogue for a smaller patient group, the workload caused by the implementation of the solution will decrease.

eDialogue was a solution where both patients and HCPs across settings could communicate freely in the postdischarge period. However, the primary communication in eDialogue was between the patient and HCPs at the hospital. Municipal physiotherapists used eDialogue more indirectly as a way to keep up to date with the patients' progress. As such, findings revealed how physiotherapists in the municipality and patients together would formulate questions to send to the hospital staff. Taking into account this shared use of eDialogue, usage data defining the proportion of messages sent between patients and HCPs and between HCPs across settings would not be representative of their actual use. Moreover, HCPs adapted eDialogue to their contexts and developed individual strategies for providing timely answers. Some strategies were developed because the technical

XSL•FO
RenderX

solution lacked better adaptation to the context, for example, an improved notification system, whereas other strategies were based on individual preferences in handling digital communication. All cases emphasize the importance of uncovering the HCPs' context and needs and ensuring that new technology supports them in their work processes so that inappropriate use of new solutions does not end up adding new workarounds and thus hindering the optimal outcome of the technology.

Limitations

This study was inspired by the CFIR to guide data collection and analysis [17,20]. The systematic identification and mapping of what was perceived as important to HCPs to the CFIR domains and constructs was helpful in providing an overview of the multifaceted and conflicting attitudes and experiences of eDialogue. However, we did not apply the CFIR as exhaustively as recommended [17,18], and we may thereby have missed important aspects that could have emerged. Using an inductive-deductive approach in data analysis, however, allowed us to still be explorative, which suited the early phase of the intervention described in this study.

In phase 1, we included a wide range of HCPs involved in the patients' trajectory and communicative circles after surgery and discharge to shed light on their perspectives on current communication pathways. Including HCPs from different settings was a strength to this study, however, the small subgroups of HCPs from the same setting may jeopardize data saturation [15]. However, the theme of the interviews, exclusively focusing on communication, is narrow and may thereby outweigh this issue. For preintervention interviews, data saturation was reached; however, it can be discussed whether data saturation was reached fully for the interviews in phase 2. Observations of HCPs' use of eDialogue, including technical or collaborative issues that were encountered during use, accounted for this and were included in the data analysis for phase 2.

Furthermore, we could have included management and decision makers in the focus group to gain a deeper understanding of the political and managerial context of the use of eDialogue across sectors. However, this was not attempted in this study as we wished to focus on the end users' perspectives.

Our findings derive from a single hospital and a municipal region in Denmark. Therefore, they may not reflect the experiences of HCPs from other parts of the country, where different digital communication solutions have been implemented. Only 1 nurse participated in communication in eDialogue, and thus the experiences of this group of HCPs are not reflected in our findings. Unfortunately, at the start of this study, the coronavirus outbreak was at its peak, and many nurses from the hospital were reassigned to newly opened COVID-19 departments. At the same time, there was a trade union strike among nurses in Denmark, which resulted in the cessation of work for a period of time for many nurses from the municipality. These circumstances put greater work pressure on the nurses, and we continued the study without their active involvement in the dialogues.

Last but not least, some of the research team members behind this study are clinicians and were involved in the decision to test eDialogue. We have tried to overcome this issue by including research team members with little knowledge of the patients and processes in orthopedic surgery. Thus, 2 independent researchers coded and condensed data (LWHJ and REKL) in close discussion with BD, where REKL did not have preliminary knowledge of the context.

Conclusions

HCPs describe current communication pathways as complicated. Phone calls are disruptive to work processes, and the lack of direct communication modalities between patients and HCPs across settings in the postoperative period makes patients become messengers of information between HCPs. To overcome these challenges, HCPs use off-the-shelf digital communication solutions as a workaround; however, use is neither standardized nor data secure. HCPs were open to using eDialogue, although they had reservations, which were partly confirmed and unconfirmed in their subsequent use of eDialogue. Especially, concerns regarding resource consumption were highlighted, and HCPs suggested the solution is particularly valuable in complex and prolonged treatments. The use of eDialogue offers a potentially valuable strategy for future integration of communication across health care settings, breaking down existing silos and taking into account the whole care team and the patient. This study provides knowledge for future strategies for implementing such solutions in orthopedic surgery and other clinical domains.

Acknowledgments

We would like to thank all patients and parents, health care providers, and management staff from departments across hospitals and municipalities in Northern Jutland, Denmark, who agreed to participate in this study. We would also like to thank the project nurse, who was an observer in focus groups, and the students from Aalborg University, who contributed to transcribing large amounts of interview data. This study was partly funded by the Research Fund of Danish Health Region North (journal number 2021-0019).

Conflicts of Interest

None declared.

References

XSL•FO
RenderX

1. Herndon JH, Pollick KJ. Continuing concerns, new challenges, and next steps in physician-patient communication. J Bone Joint Surg Am 2002;84(2):309-315. [doi: 10.2106/00004623-200202000-00019] [Medline: 11861738]
2. Naylor MD, Aiken LH, Kurtzman ET, Olds DM, Hirschman KB. The care span: the importance of transitional care in achieving health reform. Health Aff (Millwood) 2011;30(4):746-754. [doi: 10.1377/hlthaff.2011.0041] [Medline: 21471497]
3. Forster AJ, Murff HJ, Peterson JF, Gandhi TK, Bates DW. The incidence and severity of adverse events affecting patients after discharge from the hospital. Ann Intern Med 2003;138(3):161-167. [doi: 10.7326/0003-4819-138-3-200302040-00007] [Medline: 12558354]
4. Oslock WM, Satiani B, Way DP, Tamer RM, Maurer J, Hawley JD, et al. A contemporary reassessment of the US surgical workforce through 2050 predicts continued shortages and increased productivity demands. Am J Surg 2022;223(1):28-35. [doi: 10.1016/j.amjsurg.2021.07.033] [Medline: 34376275]
5. Wainwright T, Middleton R. An orthopaedic enhanced recovery pathway. Curr Anaesth Crit Care 2010;21(3):114-120 [FREE Full text] [doi: 10.1016/j.cacc.2010.01.003]
6. Pereno A, Eriksson D. A multi-stakeholder perspective on sustainable healthcare: from 2030 onwards. Futures 2020;122:102605 [FREE Full text] [doi: 10.1016/j.futures.2020.102605] [Medline: 32834076]
7. Danbjørg DB, Wagner L, Kristensen BR, Clemensen J. Intervention among new parents followed up by an interview study exploring their experiences of telemedicine after early postnatal discharge. Midwifery 2015;31(6):574-581. [doi: 10.1016/j.midw.2015.02.007] [Medline: 25765743]
8. Holm KG, Brødsgaard A, Zachariassen G, Smith AC, Clemensen J. Parent perspectives of neonatal tele-homecare: a qualitative study. J Telemed Telecare 2019;25(4):221-229. [doi: 10.1177/1357633X18765059] [Medline: 29792079]
9. Cronin RM, Davis SE, Shenson JA, Chen Q, Rosenbloom ST, Jackson GP. Growth of secure messaging through a patient portal as a form of outpatient interaction across clinical specialties. Appl Clin Inform 2015;6(2):288-304 [FREE Full text] [doi: 10.4338/ACI-2014-12-RA-0117] [Medline: 26171076]
10. Baer D. Patient-physician e-mail communication: the kaiser permanente experience. J Oncol Pract 2011;7(4):230-233 [FREE Full text] [doi: 10.1200/JOP.2011.000323] [Medline: 22043186]
11. Healthcare in Denmark: An Overview. Ministry of Health. 2017. URL: https://sum.dk/Media/C/A/Healthcare-in%20denmark%20an%20overview%20english-V16-dec.pdf [accessed 2024-02-14]
12. Murray E, Hekler EB, Andersson G, Collins LM, Doherty A, Hollis C, et al. Evaluating digital health interventions: key questions and approaches. Am J Prev Med 2016;51(5):843-851 [FREE Full text] [doi: 10.1016/j.amepre.2016.06.008] [Medline: 27745684]
13. DeWalt KM, DeWalt BR. Participant Observation: A Guide for Fieldworkers, Second Edition. Lanham, Md: AltaMira Press; 2011.
14. Morgan DL. Focus Groups as Qualitative Research, Second Edition. Thousand Oaks, CA: SAGE Publications; 1997.
15. Brinkmann S, Kvale S. Interviews: Learning the Craft of Qualitative Research Interviewing, Third Edition. København: SAGE Publications; 2015.
16. Tong A, Sainsbury P, Craig J. Consolidated Criteria for Reporting Qualitative Research (COREQ): a 32-item checklist for interviews and focus groups. Int J Qual Health Care 2007;19(6):349-357 [FREE Full text] [doi: 10.1093/intqhc/mzm042] [Medline: 17872937]
17. Damschroder LJ, Aron DC, Keith RE, Kirsh SR, Alexander JA, Lowery JC. Fostering implementation of health services research findings into practice: a consolidated framework for advancing implementation science. Implement Sci 2009;4:50 [FREE Full text] [doi: 10.1186/1748-5908-4-50] [Medline: 19664226]
18. Kirk MA, Kelley C, Yankey N, Birken SA, Abadie B, Damschroder L. A systematic review of the use of the consolidated framework for implementation research. Implement Sci 2016;11:72 [FREE Full text] [doi: 10.1186/s13012-016-0437-z] [Medline: 27189233]
19. First Agenda. LetDialog—Sikker SMS. URL: https://letdialog.dk/ [accessed 2023-02-02]
20. The CFIR interview guide tool. CFIR Research Team-Center for Clinical Management Research. URL: https://cfirguide.org/tools/ [accessed 2023-02-02]
21. NVivo Qualitative data analysis software. QSR International. 1999. URL: https://qsrinternational.com/nvivo/nvivo-products/ [accessed 2023-02-02]
22. Lindkvist RM, Sjöström-Strand A, Landgren K, Johnsson BA, Stenström P, Hallström IK. "In a way we took the hospital home"-a descriptive mixed-methods study of parents' usage and experiences of eHealth for self-management after hospital discharge due to pediatric surgery or preterm birth. Int J Environ Res Public Health 2021;18(12):6480 [FREE Full text] [doi: 10.3390/ijerph18126480] [Medline: 34203985]
23. DeAntonio JH, Kang HS, Cockrell HC, Rothstein W, Oiticica C, Lanning DA. Utilization of a handheld telemedicine device in postoperative pediatric surgical care. J Pediatr Surg 2019;54(5):1005-1008. [doi: 10.1016/j.jpedsurg.2019.01.032] [Medline: 30782441]
24. Gund A, Sjöqvist BA, Wigert H, Hentz E, Lindecrantz K, Bry K. A randomized controlled study about the use of eHealth in the home health care of premature infants. BMC Med Inform Decis Mak 2013;13:22 [FREE Full text] [doi: 10.1186/1472-6947-13-22] [Medline: 23394465]

XSL•FO
RenderX

25. Jensen B, Vardinghus-Nielsen H, Mills EHA, Møller AL, Gnesin F, Zylyftari N, et al. "Like a rainy weather inside of me": qualitative content analysis of telephone consultations concerning back pain preceding out-of-hospital cardiac arrest. Int Emerg Nurs 2022;64:101200 [FREE Full text] [doi: 10.1016/j.ienj.2022.101200] [Medline: 35926318]
26. Ekstedt M, Børøsund E, Svenningsen IK, Ruland CM. Reducing errors through a web-based self-management support system. Stud Health Technol Inform 2014;201:328-334. [Medline: 24943563]
27. Voruganti T, Grunfeld E, Jamieson T, Kurahashi AM, Lokuge B, Krzyzanowska MK, et al. My team of care study: a pilot randomized controlled trial of a web-based communication tool for collaborative care in patients with advanced cancer. J Med Internet Res 2017;19(7):e219 [FREE Full text] [doi: 10.2196/jmir.7421] [Medline: 28720558]
28. Husain A, Cohen E, Dubrowski R, Jamieson T, Kurahashi AM, Lokuge B, et al. A clinical communication tool (Loop) for team-based care in pediatric and adult care settings: hybrid mixed methods implementation study. J Med Internet Res 2021;23(3):e25505 [FREE Full text] [doi: 10.2196/25505] [Medline: 33656445]
29. Gulmans J, Vollenbroek-Hutten M, van Gemert-Pijnen L, van Harten W. A web-based communication system for integrated care in cerebral palsy: experienced contribution to parent-professional communication. Int J Integr Care 2012;12:e9 [FREE Full text] [doi: 10.5334/ijic.672] [Medline: 22977435]
30. Balakrishnan AS, Palmer NR, Odisho AY. "Can you hear me now?": postoperative patient-initiated communication with providers. Am J Manag Care 2020;26(10):e333-e341 [FREE Full text] [doi: 10.37765/ajmc.2020.88507] [Medline: 33094946]

Abbreviations

CFIR: Consolidated Framework for Implementation Research
COREQ: Consolidated Criteria for Reporting Qualitative Research
HCP: health care professional

Edited by G Eysenbach; submitted 09.10.23; peer-reviewed by A Mahnke, T Tang; comments to author 02.11.23; revised version received 21.11.23; accepted 31.01.24; published 08.03.24.

Please cite as:
Jensen LWH, Rahbek O, Lauritsen REK, Kold S, Dinesen B
Health Care Professionals' Perspectives Before and After Use of eDialogue for Team-Based Digital Communication Across Settings: Qualitative Study
JMIR Hum Factors 2024;11:e53391
URL: https://humanfactors.jmir.org/2024/1/e53391
doi:10.2196/53391
PMID:38457798

©Lili Worre Høpfner Jensen, Ole Rahbek, Rikke Emilie Kildahl Lauritsen, Søren Kold, Birthe Dinesen. Originally published in JMIR Human Factors (https://humanfactors.jmir.org), 08.03.2024. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on https://humanfactors.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Exploring the Role of Complexity in Health Care Technology Bottom-Up Innovations: Multiple-Case Study Using the Nonadoption, Abandonment, Scale-Up, Spread, and Sustainability Complexity Assessment Tool

Ulla Hellstrand Tang[1,2*], PhD; Frida Smith[3,4*], PhD; Ulla Leyla Karilampi[5], PhD; Andreas Gremyr[5,6], PhD

[1]Department of Prosthetics and Orthotics, Sahlgrenska University Hospital, Gothenburg, Sweden

[2]Department of Orthopaedics, Institute of Clinical Sciences, Sahlgrenska Academy at University of Gothenburg, Gothenburg, Sweden

[3]Regional Cancer Centre West, Gothenburg, Sweden

[4]Department of Technology Management and Economics, Collaborative Plattform for Healthcare Improvement, Chalmers University of Technology, Gothenburg, Sweden

[5]Department of Schizophrenia Spectrum Disorders, Sahlgrenska University Hospital, Gothenburg, Sweden

[6]Jönköping Academy for Improvement of Health and Welfare, School of Health and Welfare, Jönköping University, Jönköping, Sweden

[*]these authors contributed equally

Corresponding Author:
Ulla Hellstrand Tang, PhD
Department of Prosthetics and Orthotics, Sahlgrenska University Hospital
Falkenbergsgatan 3
Gothenburg, SE-412 85
Sweden
Phone: 46 706397913
Fax: 46 31408162
Email: ulla.tang@vgregion.se

Abstract

Background: New digital technology presents new challenges to health care on multiple levels. There are calls for further research that considers the complex factors related to digital innovations in complex health care settings to bridge the gap when moving from linear, logistic research to embracing and testing the concept of complexity. The nonadoption, abandonment, scale-up, spread, and sustainability (NASSS) framework was developed to help study complexity in digital innovations.

Objective: This study aims to investigate the role of complexity in the development and deployment of innovations by retrospectively assessing challenges to 4 digital health care innovations initiated from the bottom up.

Methods: A multicase retrospective, deductive, and explorative analysis using the NASSS complexity assessment tool LONG was conducted. In total, 4 bottom-up innovations developed in Region Västra Götaland in Sweden were explored and compared to identify unique and shared complexity-related challenges.

Results: The analysis resulted in joint insights and individual learning. Overall, the complexity was mostly found outside the actual innovation; more specifically, it related to the organization's readiness to integrate new innovations, how to manage and maintain innovations, and how to finance them. The NASSS framework sheds light on various perspectives that can either facilitate or hinder the adoption, scale-up, and spread of technological innovations. In the domain of condition or diagnosis, a well-informed understanding of the complexity related to the condition or illness (diabetes, cancer, bipolar disorders, and schizophrenia disorders) is of great importance for the innovation. The value proposition needs to be clearly described early to enable an understanding of costs and outcomes. The questions in the NASSS complexity assessment tool LONG were sometimes difficult to comprehend, not only from a language perspective but also due to a lack of understanding of the surrounding organization's system and its setting.

Conclusions: Even when bottom-up innovations arise within the same support organization, the complexity can vary based on the developmental phase and the unique characteristics of each project. Identifying, defining, and understanding complexity may not solve the issues but substantially improves the prospects for successful deployment. Successful innovation within complex organizations necessitates an adaptive leadership and structures to surmount cultural resistance and organizational impediments.

XSL•FO
RenderX

A rigid, linear, and stepwise approach risks disregarding interconnected variables and dependencies, leading to suboptimal outcomes. Success lies in embracing the complexity with its uncertainty, nurturing creativity, and adopting a nonlinear methodology that accommodates the iterative nature of innovation processes within complex organizations.

(JMIR Hum Factors 2024;11:e50889) doi:10.2196/50889

KEYWORDS

digital; bottom-up innovation; complexity; eHealth; health care; nonadoption, abandonment, scale-up, spread, and sustainability complexity assessment tool; NASSS-CAT; mobile phone

Introduction

Why Is it so Difficult to Develop and Spread New Innovative Technologies in Health Care?

There has been an increasing focus on innovation and the role of new technologies (eg, electronic health records, smartphones, and health applications) in health care. However, developing new technologies comes with significant challenges. Studies show that technology projects in health care, particularly large and complex projects, have a high rate of failure and seldom produce the anticipated results [1-5]. Bottom-up innovations in health care are innovations for service delivery that have been developed "from the ground up," often focusing on preventive patient-centered care, typically driven forward by small interdisciplinary groups of professionals and patients [6-9]. As a result, they may not be captured by existing metrics, thus being "invisible" to senior management and policy makers [10].

The challenges when it comes to developing and making use of innovations, such as spreading or implementing new ways of working [11], have been described by many, often as a "knowledge translation or production problem" [12]. Braithwaite et al [13] compared the traditionally dominant linear and causal thinking that characterizes early implementation science and the evidence-based medicine paradigm with features such as those in systems thinking. The linear approach applies simple, orderly processes with cumulative sequences of stages to produce results building on a knowledge of the way things work, making use of predictable relationships between causes and effects. This has helped generate many successes in the past, not least in health care, but it tends to increase rigidity and fail when applied in more complex and messy systems where things change dynamically and are therefore unpredictable [13]. Instead, systems thinking and the related complexity science recognize system characteristics in building an understanding of how best to move forward.

To drive change in a predictable "simple" system where causes and effects are known, a linear, stepwise approach has a greater chance of success. However, complex systems are not only dynamic but also are often described as adaptive (as in complex adaptive systems) in that they are constituted of agents and artifacts that communicate and learn from each other and the surrounding environment, creating opportunities to learn from experience, self-organize, and evolve, making them less predictable systems [14].

Even though the linear approach has previously dominated implementation and development initiatives in health care, many researchers point to the necessity to apply systems thinking and complexity science when developing health care through the innovative use of new technologies, as exemplified in the study by Greenhalgh and Papoutsi [2]. As the concept of stepwise, linear cause and effect is not sufficient when studying complex systems that evolve in ways that are impossible to predict, it is relevant to use the knowledge of complex systems when understanding and studying health services [15]. Complex systems are defined by (1) intricate intertwined processes, (2) interconnectivity between systems, (3) interconnectivity between levels within systems, and (4) interconnectivity between actors and elements, giving complex systems different properties from those of less complex systems [1]. In short, a complex system does not work linearly but dynamically, with fundamentally different logics [16], and needs to be addressed and understood accordingly during innovation and implementation. If not, there is a risk that new technology and innovations will further increase the complexity rather than actually supporting the needs and demands for an improved health care system [17].

The Challenges

A total of 4 bottom-up innovators found that there was a need to gain insights into the complexity involved in developing and executing bottom-up innovations in a complex health care organization. All 4 innovators had met with hindrances preventing them from moving forward with their innovations. It was necessary to pause and retrospectively try to comprehend the underlying reasons for the stagnation in the 4 cases in question.

Project representatives, all health care professionals, joined forces to identify challenges by assessing project complexity to increase an understanding of the role of complexity and find ways to explore and assess it. As they all worked within the same regional system, it was crucial to involve regional stakeholders (support functions) during the learning process.

The aim of this study was to investigate the role of complexity in the development and deployment of innovations by retrospectively assessing the challenges to 4 digital health care innovations initiated from the bottom up.

Methods

This section describes the theoretical framework that underlies our methodological approach, the settings, and the 4 cases under study, as well as the procedure.

Theoretical Framework

An impressive amount of knowledge related to the diffusion of innovations and their implementation in health care by the start of the new millennium is summarized in the extensive review

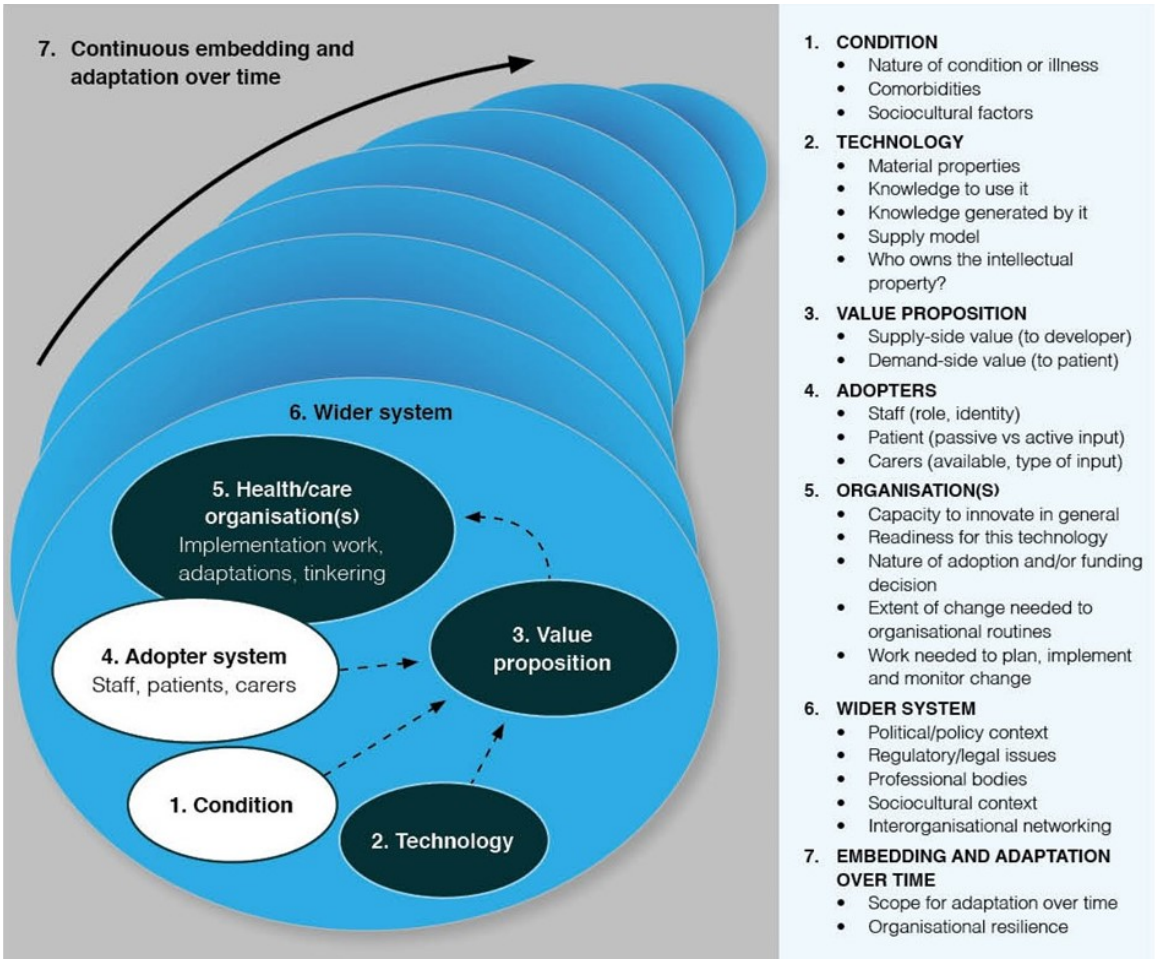
XSL•FO
RenderX

by Greenhalgh et al [11] from 2004. It builds partly on the ideas by Rogers [18] that innovations have characteristics that will affect their diffusion, as well as affect other domains (eg, the readiness of the system for change, the implementation process, the adopter, and the external wider [sociopolitical] context). As innovations in health care were increasingly associated with new technologies, a new review was conducted by Robert et al [19] in 2010, adding more recent data and focusing on the adoption and assimilation of new technologies into health care.

This, along with the high failure rate of health care technology innovation projects, inspired Greenhalgh and colleagues to deepen their knowledge of the diffusion of innovations, with an emphasis on health technology projects. Building on previous work, reviewing the literature, and using empirical studies of technology implementation, they elaborated on and explained domains of importance. This resulted in the nonadoption, abandonment, scale-up, spread, and sustainability (NASSS) framework (Figure 1 [20], published under Creative Commons Attribution 4.0 International License, CC BY).

The NASSS framework was developed into a complexity assessment tool (NASSS-complexity assessment tool [NASSS-CAT]) [20] to help assess the complexity of health technology projects before, during, or after they were finished.

Figure 1. The nonadoption, abandonment, scale-up, spread, and sustainability complexity assessment tool with its 7 domains.

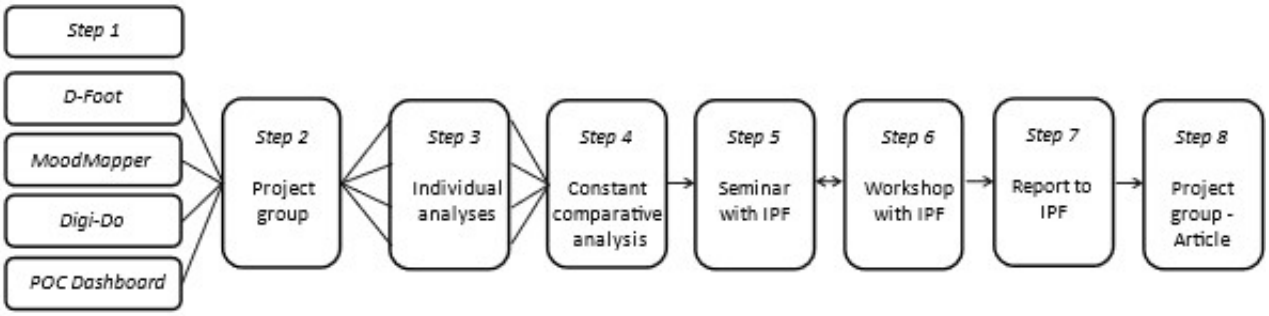


Design and Methodological Approach

A multicase retrospective, deductive, and explorative analysis using the NASSS-CAT LONG was conducted [15,21]. The process of analysis is shown in Figure 2 and is described at the end of the *Methods* section. The complexities of 4 bottom-up innovations developed in Region Västra Götaland (VGR) in Sweden were explored. The NASSS-CAT LONG consists of 2 parts divided into 7 domains (Figure 1). First, one is asked to describe the project and its potential messiness in their own words. Writing this narrative can help surface interdependencies and tricky issues of the project, hence revealing complexity. Second, one answers the questions related to the domain to help them estimate key areas of complexity. One can define whether the question is complex or not complex, whether they do not know, or whether it is not applicable. The total score of orange boxes ticked tells one how complex a certain domain is for their project. In part 2, one is guided through prompts to help them plan for and manage complexity by *reducing* it where possible and *responding* to it if or where it cannot be reduced. The questions can be answered by different people who will provide the needed insights into the domain and the project under evaluation.

XSL•FO
RenderX

Figure 2. Flowchart illustrating the study process when exploring the role of complexity in health care technology projects. This figure presents the different steps (1-8) in the study. Details can be found in the main text. IPF: Innovation Platform.



Setting

The 4 bottom-up innovators were from different parts of the organization in the VGR [22]. All of them had unique experiences of their own departments of the VGR and the surrounding supporting systems, such as the IT departments, purchasing departments, and legal offices. Although all 4 worked in specialized care, the care flow for each of the relevant diagnoses spanned specialized care, primary care, and municipal care.

The VGR is a region on the western coast of Sweden with a total population of 1.8 million. For several years, there has been a push for increasing the development of innovative solutions to the challenges faced by, for example, the public health care system. This has been implemented through the formation of an Innovation Platform (IPF) that develops processes and supportive structures as well as approving funds to help innovative ideas thrive. The mission of the IPF is to contribute to a sustainable innovation system that promotes innovation in health care and ensures that collaboration between academia and business fulfills the needs of patients within the health care system.

In total, 4 bottom-up innovators in the VGR pioneering eHealth and clinical research united in 2019 after realizing that there were barriers to their separate innovations due to an organizational lack of an innovation framework and to inexperience in developing, testing, implementing, and maintaining digital innovations. One author, familiar with the NASSS framework, encouraged the others to explore complexity in innovation. Together, they adapted the NASSS-CAT to identify unique and shared complexities in their innovations. Concurrently, at the same network meeting, representatives from the IPF wanted to be a part of the study, exchanging insights on how to identify and manage complexities in the innovation process.

Ethical Considerations

Because no personal data were collected, no ethics approval was needed. All participants agreed, orally, to take part. No sensitive personal information was collected, and no patients participated in the workshops. The IPF, the regional support resource for innovations in health care, read and commented on the Swedish report before publication.

The 4 Bottom-Up Innovations

Overview

Each innovation is presented in the following sections and in Multimedia Appendix 1 [15,20,23-38]. The cases are heterogeneous with regard to intended user, phase in the innovation process, and place of implementation (locally, regionally, and nationally). Despite their differences, they all existed in the same environment, framed by the regulations of the VGR and its support system for innovation.

Case 1: The D-Foot

The lifetime risk of developing diabetic foot ulcers (DFUs) is as high as 34% for patients with diabetes [23]. It is a burden for the patient and for health care with regard to costs. With prompt prevention, the prevalence of DFUs can be halved [24].

The D-Foot is a digital decision support system designed for preventing DFUs. It conducts early screening and provides treatment recommendations based on a risk grade (ranging from 1=no risk to 4=ongoing foot ulcers) [25,26]. The risk grade is automatically generated through a series of structured foot assessments and patient surveys [27]. A printable report of foot assessments, risk grade, and recommendations is generated. The innovation's reliability and usability have previously been reported and assessed as good [27,28]. The innovator's intention was that the D-Foot would serve as a tool in the national effort to implement a person-centered and seamless care chain for preventing foot ulcers in people living with diabetes [25,26].

The seamless care chain consists of (1) an annual foot examination, (2) a podiatry intervention, (3) the provision of appropriate footwear for at-risk patients, and (4) treatment in a multidisciplinary team for patients with active DFUs [25,39]. The D-Foot was developed as an easy-to-use digital tool to support foot examinations for individuals diagnosed with diabetes, primarily targeting prosthetic and orthotic specialist care [27,28,40]. The goal was to implement the D-Foot nationally, expecting early prevention of DFUs, improved quality of life for affected individuals [29,41], and reduced health care costs [42].

Version 1.0 of the D-Foot software was developed from 2011 to 2016 by an expert group comprising certified prosthetists and orthotists, patient representatives, and orthopedic surgeons in the VGR [27]. Initially, it underwent regional testing with positive results. Thereafter, continuous improvements have been made based on users' comments [28]. Not yet executed is the

XSL•FO
RenderX

request from users for integration between the D-Foot and the major medical record system [28].

Case 2: The MoodMapper

Bipolar disorder, often diagnosed in early adulthood, typically necessitates lifelong treatment. It leads to undesirable mood swings affecting daily functioning. Mood episodes vary from extreme "highs" (manic episodes) to severe "lows" (depressive episodes) lasting for days or weeks. Even with proper treatment, mood fluctuations can occur. Collaborative communication between patients and health care providers enhances treatment effectiveness. Moreover, the early detection of behavior changes is of the utmost importance in the successful treatment of bipolar disorder.

The aim of this innovation was to determine whether smartphone use data are a reliable source for studying changes in the digital behavioral patterns of individuals with bipolar disorder by exploring correlations between different parameters of smartphone use data.

The MoodMapper is a mobile app that, through real-time data collection, can provide valuable insights into a patient's smartphone use. The ambition was, after pilot-testing, to study and evaluate the connection between mobile-generated passive data and documented changes in the patient's mental well-being, with the goal of making it easier for both patients and health care professionals to monitor the progression of the patient's condition and make decisions regarding prevention and care.

Case 3: The Digi-Do

Radiation therapy (RT) is a common treatment after breast cancer surgery. The high-technology environment and unfamiliar nature of RT can affect the patient's experience of the treatment. Misconceptions or a lack of knowledge related to RT processes can increase levels of anxiety and enhance feelings of being unprepared at the beginning of the treatment. Moreover, the waiting time is often fairly long. Cancer care involves several, often independent clinics. Even if the clinical pathway is clearly described, transitions and information exchange can be problematic. RT is only provided at the university hospital in the region, with long distances and long waiting times for many patients.

The Digi-Do tool consists of two separate mobile apps: (1) an app providing a guided digital tour of the RT department, where the patient can familiarize themselves with the department by using virtual reality glasses; and (2) an app with additional information, including questions and answers, practical information, and short animated films about the RT process. The design of both apps was developed in a co-design process with patients and staff [43]. The primary aim of the researcher or innovator was to evaluate whether a digital information tool with virtual reality technology and preparatory information was able to reduce distress and enhance the self-efficacy and health literacy of patients with breast cancer before, during, and after RT. A secondary aim was to explore whether the digital information tool increased patient flow while maintaining or improving the quality of care [44].

Case 4: A Point-of-Care Dashboard for Schizophrenia Care (the PoC Dashboard)

The Department of Schizophrenia Spectrum Disorders at Sahlgrenska University Hospital delivers specialized care for people with psychotic disorders in the metropolitan Gothenburg area (with a population of approximately 600,000 people) in Sweden. Schizophrenia is the most common diagnosis among the approximately 3000 patients who receive care at the department's 7 outpatient units. Approximately 20% of these patients also need acute inpatient care at 1 of the department's 4 wards each year.

With the aim of supporting patient coproduction of health, a digital dashboard was developed to be jointly reviewed at the point of care by patients and case managers and psychiatrists to support evaluation and planning, outcome questionnaires, and patients' care plans [45]. The dashboard was developed between 2016 and 2018 and was piloted at 2 outpatient units with approximately 400 patients for 18 months. The dashboard is one of several connected applications and displays for visualizing data fed by multiple systems to support, for example, planning, management, and triage and includes a unit-level overview of quality indicators to identify patients at risk. The dashboard project also served as a case in the development of the NASSS-CAT [20].

Study Procedure

This study followed an iterative process, including analyses, discussions, and seminars (Figure 2).

As we used the NASSS-CAT, the analysis was deemed to be of an exploratory, deductive nature. First, an individual assessment of each case was made, and then the 4 cases were compared to find similarities. However, while using the NASSS-CAT in each of the 4 cases, the innovators had discussions about how to interpret the questions in the 7 domains. An additional method, namely, constant comparative analysis (CCA), was chosen as it is appropriate in collaborative projects to facilitate and identify agreements and disagreements [46,47] (Multimedia Appendix 2). After agreeing on how to interpret the questions, members of the IPF were invited to complete the analysis in a workshop.

In total, 4 bottom-up innovators had individually experienced complexities during their respective innovation processes from 2010 to 2019.

1. The 4 innovators got together and started the study in January 2020 by learning how to use the NASSS framework based on the work by Greenhalgh et al [15]. Support was available as one of the authors had been involved in the development of the NASSS-CAT [20].
2. The innovators identified complexities in their individual projects using the NASSS-CAT [20] in 2020.
3. During >30 one-hour meetings using the NASSS-CAT and taking minutes, the innovators identified, compared, and discussed similarities and differences regarding complexities in their respective innovations. A CCA was included in the process and is described in Multimedia Appendix 2.

XSL•FO
RenderX

4. A seminar was held in April 2020 with one of the bottom-up innovators and the IPF presenting the concept of complexity and the NASSS framework.
5. In October 2020, another seminar was held with the 4 innovators and staff members from the IPF to discuss the domains of the NASSS-CAT, illustrated by examples and findings from the assessment of the 4 bottom-up innovations. The participants from the IPF discussed and reflected on experiences of complexities. The seminar was recorded and summarized in a report in collaboration with a representative from the IPF [48].
6. The authors were commissioned to write a report (in Swedish) for the IPF exploring and summarizing the NASSS framework and complexity with examples from the bottom-up innovations [48].
7. The insights gained into the role of complexities from the entire aforementioned process were discussed and summarized and are presented in this paper.

Results

In this retrospective exploration of the role of complexity in 4 bottom-up health care innovations in a Swedish region, both similarities and differences emerged among the 4 cases when using the NASSS-CAT (Multimedia Appendix 3). The findings for each domain are described in the following sections.

Complexity Domain 1: The Illness or Condition

This domain has no or low complexity when the illness is well known and an assessment can result in a well-defined diagnosis and when there is, furthermore, knowledge and know-how regarding how to treat the condition successfully. Complexity can be related to conditions with less known causes, a high prevalence of multimorbidity, and challenging sociocultural factors (eg, language barriers). In our study, the cases that addressed mental illness (PoC Dashboard and MoodMapper) and diabetes (D-Foot) had more complexity related to the actual illnesses than the case aiming to prepare women before RT for breast cancer (Digi-Do) [49,50]. Both bipolar disorder and schizophrenia are strongly connected with comorbidity and lifestyle-related conditions, and even though national guidelines exist, there is no simple pathway to treat those conditions. The third case (D-Foot) involved diabetes, also an illness defined as complex due to the patient being treated in various institutions and with several lifestyle factors influencing the outcome of the treatment [39].

Complexity Domain 2: The Technology (or Other Innovations)

An innovation is less complex if it is well known, ready, and easy to use and has a clear supply model, well-defined ownership regarding its intellectual properties, and low or no dependency on other systems. For the actual technologies under study, similarities regarding complexities revolved around interdependencies with other IT systems, ranging from local to regional and even national systems. Even if the technology already existed (D-Foot) or if new software was developed to create a better overview of data in several existing systems (PoC Dashboard), it was difficult to develop the innovation so that it enabled adoption beyond the local settings. Regulations regarding software used as a medical device [30] sometimes prevailed over the simple adaptations for different target groups. “Fireproof” walls exist between organizations (eg, municipal care, primary care, and specialist care), and different versions of the regional information systems, being related to ownership, budget, and management, make it less clear whether and how new technologies can be bought, adapted, and used in a local setting. In contrast, the Digi-Do app does not require any interaction with existing IT systems and was not deemed to be a medical technical device. The MoodMapper app, on the other hand, is complex as it aims to interact with both patients and health care staff, requiring interaction with medical electronic health records as well as ensuring a very high level of security to safeguard the patients’ integrity [51]. The need for supply chains included both purchasing and procurement and clinical implementation, the latter involving questions regarding intellectual properties (is the owner the bottom-up innovator or is it the region?), ownership management (which regulation steers the region when managing a medical device owned by the region?), and updates and maintenance of the eHealth tools (which department in the VGR is responsible for updates and maintenance of the innovations?). All the cases had run into or expected to run into severe complexity when planning to launch their innovations. It was clear that complexity regarding supply chains had not been considered by either the innovator or the VGR when intending to expand from a local level to a regional or national one. An example of the complex challenges related to the spread and maintenance of one of the eHealth tools, the D-Foot, is presented in the following paragraphs.

Regulations regarding funding and ownership made it difficult to implement a supply chain outside the local region as each of Sweden’s 21 regions has its own procurement processes. Since the start, the D-Foot project had been aiming for national spread. In 2017, the IPF approved funding with the aim of testing, Conformité Européenne (CE) marking, and thereafter implementing the D-Foot first in the VGR at the department of prosthetics and orthotics and then nationally. At the same time, several departments of prosthetics and orthotics in other regions were interested in using the D-Foot as soon as the CE marking was finalized. However, in June 2017, an official at the VGR decided that the region was only able to allow the D-Foot to be used within the region (Article 5.5 in the Regulation [European Union] 2017/745) [30]. Following this decision, national spread was impossible. The bottom-up innovator continued to have a dialogue with the IPF seeking a solution for national spread. In 2020, an opportunity for national spread arose by registering the D-Foot as a national medical information system (NMI) at the Medical Products Agency. An NMI is an information system developed for joint use at nationwide, regional, or municipal level in Sweden.

Thus, the D-Foot transitioned from being a “self-manufactured medical device” to becoming an NMI registered with the Swedish Medical Products Agency in 2020. However, the NMI registration was withdrawn by the VGR in 2021 due to new regulations from the Swedish Medical Products Agency [52]. The D-Foot remains a separate software program not integrated into the standard medical record system in the VGR. As a result,

XSL•FO
RenderX

one option remained for national spread, namely, to CE mark the D-Foot, a procedure that was not as yet allowed or tested in the region.

Complexity Domain 3: The Value Proposition (Costs and Benefits of the Technology)

Complexity in this domain arises when determining the value provided by the innovation to developers, users (patients, staff, and health care systems), and the broader health care ecosystem. Despite their origin as bottom-up innovations aimed at improving care, the complexities of demonstrating supply-side value in terms of business models and monetary benefits were challenging. Key questions included defining improved value, decision-making processes, the inclusion of nonmonetary values, and the extent of evaluation required: what is regarded as improved value? Who decides? Is it only monetary or other types of value as well? How much does the innovation need to be evaluated and how?

The Digi-Do has a defined regional vision and has faced challenges in quantifying value, especially regarding soft values such as reduced distress and increased health literacy and self-efficacy. The Digi-Do aims to optimize the use of waiting time before RT, adding value by reducing waiting times and queues. Furthermore, the innovation aims to create value for patients by delivering information in a novel, accessible format, potentially improving health literacy even for those with language difficulties or cognitive impairments. It also extends benefits to the patient's social network, enhancing support and knowledge and reducing distress among family and friends affected by the patient's cancer diagnosis. By using the often idle waiting time for meaningful preparation, the innovation may foster a sense of control and inclusion, diminishing distress and worry. Well-prepared patients may navigate the system more efficiently, potentially reducing waiting times for information dissemination.

The evaluation of the outcomes in this specific project is still ongoing through an unpublished randomized controlled trial [44], but so far, the qualitative results show a high level of acceptance of and positivity toward the tool. Nevertheless, there needs to be a discussion about how to endorse a more pragmatic evaluation of both effectiveness and process outcomes [53].

If successful, this approach could be adapted for other health care domains, although commercialization is not the project's primary goal. Measuring soft values has proved challenging as they might not directly impact traditional health care outcomes. The other cases faced similar difficulties in pinpointing the exact stages in which costs and values could be calculated.

Enhancing foot health can improve the quality of life of patients and reduce health care costs associated with treating DFUs and amputations. Objective risk assessment by using the D-Foot precedes interventions, aligning with the vision of providing equal, high-quality care to citizens. Early interventions in the prevention process (D-Foot) might require more resources within primary care but were expected to be cost-effective in the long term due to a reduction in specialist care following fewer ulcer treatments and amputations [31,42]. In terms of quality of life and cost reduction, the value proposition needs further evaluation over a longer period relying on data related to care costs for at-risk patient groups. The D-Foot database contains valuable information on risk groups and foot status, serving as a data source for audits and evaluations to optimize foot care. It could also function as a quality registry, potentially becoming the new diabetic foot register in Sweden.

The MoodMapper aims to provide a more objective risk analysis and early interventions, potentially preventing hospitalization. In the examples of innovations (MoodMapper and PoC Dashboard) designed to prevent relapses in severe mental illness by coordinating data or even by asking patients to send and react to data, the need for hospital care could be reduced. However, this area is as yet unexplored.

The value proposition of the PoC Dashboard remains uncertain. Case managers and patients find the technology useful based on preliminary data. Local testing and piloting suggest perceived effectiveness, although the degree of cost-effectiveness is still unknown. The dashboard streamlines administrative tasks for staff, offering an overview of patients' progress and risks while facilitating collaborative care planning. However, the technology's potential as a commercial product is uncertain, mainly because it is integrated with older systems. Additional uncertainties involve the IT department's role in dashboard maintenance and associated costs.

Complexity Domain 4: The Intended Adopters of the Innovation and Technology

Complexity in this domain is higher when adopting the innovation, necessitating changes in routines, roles, and identities. Innovations that support existing routines with minimal disruption are associated with lower complexity, and all 4 cases required either behavior changes by patients or modifications to work routines for health care staff. For example, the PoC Dashboard simplified patient overview and reduced administrative work for staff, thus positively impacting daily tasks [54].

However, transferring the D-Foot to primary care posed challenges as different health care professionals (podiatrists, nurses, and physicians) with varying roles and routines questioned its added value. The MoodMapper required patients to trust the handling of their behavioral data, which could be challenging for those with symptoms of paranoia.

Complexity Domain 5: The Organization Implementing the Technology

Complexity in this domain pertains to the efforts required to plan, implement, and monitor the innovation's adoption, as well as to the organization's overall capacity for innovation. Challenges included a lack of clear pathways for support, making it necessary to find the right individuals at the right levels for consultations. Different organizational levels faced varying complexities, and despite a desire for innovation, built-in regulations sometimes hindered dissemination. For instance, regulatory obstacles prevented the national spread of the D-Foot.

XSL•FO
RenderX

Complexity Domain 6: The External Context for Innovation

Complexity in this domain is influenced by the political, sociotechnical, and regulatory context, as well as by stakeholder groups and interorganizational networking. In Sweden, despite a national Vision for eHealth by 2025 [55], the existence of 21 independent regions creates complexity in decision-making for local, regional, and national development and for the implementation of digital health care innovations. For instance, there is a national initiative from the government to improve cancer care, but the regions are self-governed in terms of budget and implementation. This means that, even if the regional cancer center had a national assignment to improve cancer care generally and the RT process in the local region specifically, it has no mandate to implement the Digi-Do without the approval of the RT department at each separate regional hospital.

Furthermore, if an innovation needs to be integrated with the IT systems, such as in the other 3 cases, national initiatives can be ruled out by regional procurement, management supply chains, and European regulations regarding medical devices [30]. If regional support and the management of supply chains only permit regional use, there will be no dissemination, and thus, bottom-up innovations risk becoming only local or, at worst, experiencing the "death of innovations" after the initial project phase.

Complexity Domain 7: Emergence Over Time

Complexities were identified in all 4 cases (Multimedia Appendix 3). When summarizing the complexities from domains 1 to 6, all the authors concluded that the complexities were likely to increase in the coming 3 to 5 years, probably due to advances in technology, unexpected events such as pandemics, international conflicts, and new regulations and standards. In the coming years, a new regional medical record system, Millennium, is planned to be implemented. For small bottom-up innovators, it is not yet clear how the implementation of Millennium will affect their innovations [37].

Discussion

Principal Findings

Overview

In this study, we conducted a retrospective, deductive, and exploratory analysis of 4 cases using the NASSS-CAT LONG. We intended to explore whether there were shared or individual challenges related to bottom-up innovation projects in the same health care region. The analysis itself was complex, but it resulted in both common and individual learning, and all but one case have moved forward, partly due to new insights gained that have made progress possible. By applying the NASSS-CAT in various projects, the authors learned several lessons, the most important of which are described in the following sections. After that, we discuss and reflect on the methodology and the need or suggestions for further research. Finally, we briefly present how the cases have developed since the analysis.

The Innovation Versus the System

As proposed by Rogers [18], the properties of the innovation affect its probability of diffusion within and beyond the organization. Through this study, we have become aware of the need to understand the "system" in which we are working to develop and adopt innovations and make effective use of those innovations. The NASSS framework sheds light on various perspectives that can either facilitate or hinder the adoption, scale-up, and spread of technological innovations. Before our projects, none of us had fully considered all these perspectives. During this study, complexity was found and highlighted, involving many issues related to the organization or system rather than the specific innovation itself. Multiple regulations must be considered, and regional procurement [56], management of supply chains, and European regulations regarding medical devices [30] can hinder the spread of innovation. A lack of necessary interorganizational networking further complicates matters.

Linear Logic Versus Dynamic Complex Processes

By applying the NASSS framework, we discovered how the innovation process and the training we had all had in evidence-based medicine and the research process were geared toward a linear process rather than embracing complexity. We also discovered that the complexity was mostly found outside the actual innovation and related more to the system that the innovation was supposed to live in and to regulations and legislation. More specifically, it related to the organization's willingness to integrate new innovations and to questions regarding how to manage, maintain, and finance innovations.

Developing and deploying new bottom-up innovations in health care involves multiple logics [57]. Initially, we attempted to approach this in a traditional linear fashion with sequential steps from idea to widespread adoption. However, we quickly realized that this linear approach did not align with the reality of navigating the complexities of health care innovation. Instead of a straightforward innovation journey, it often felt like traversing a dense jungle, making it challenging, if not impossible, to gain a comprehensive overview of the landscape, identify opportunities, and predict the appropriate course of action.

Complex environments often require creative and dynamic thinking; in contrast, a linear approach may stifle the ability to respond to unexpected challenges or opportunities. Innovation is inherently uncertain and unpredictable [58]. It often involves trial and error, experimentation, and the willingness to explore unconventional ideas. A rigid stepwise approach may not accommodate the iterative and nonlinear nature of the innovation process. Complex organizations involve numerous interconnected variables and dependencies. A linear approach may overlook these interconnections, leading to suboptimal solutions or unintended consequences. Innovation often requires a holistic understanding of the organization's ecosystem. This understanding is hindered if established cultures in the complex organizations are resistant to change [57]. A nonlinear approach may face resistance from employees or departments unwilling to deviate from established norms.

XSL•FO
RenderX

Successful innovation requires addressing cultural and organizational barriers, which may not fit neatly into a linear plan. Finally, complex organizations require adaptive leadership that can navigate ambiguity and inspire a culture of continuous improvement [59]. These are important findings as innovation, particularly in the realm of new technologies, is often seen as a potential solution to address the challenges facing health care. Calls for innovation and new ways of working have come from various sources, including governments, health care organizations, and life sciences clusters. However, the high failure rate of health care technology projects suggests that there may be deficiencies in the structure, resources, and knowledge needed for success [60]. Furthermore, there is a risk of simplifying the complex innovation process by building a support system that is linear. The linear and stepwise approach (first do this, then do that) is counterproductive. While a linear approach may work in certain situations, the nature of innovation in complex organizations demands a more flexible, adaptive, and nonlinear methodology. Embracing uncertainty, fostering creativity, and adapting to change are critical elements that a rigid stepwise approach may not adequately address in the context of complex organizational innovation [58].

Value

For all 4 cases, questions arose related to value and costs. Will there be an initial or a recurrent cost for the product, or will the cost be related to a new service that entails new tasks for staff? There is an advantage in specifying both costs and values, as well as the effect of the innovation on other resources, early in the innovation process. Therefore, health-economy analyses are needed, but they are difficult to design and perform as some innovations focus on increasing soft values that are difficult to translate into monetary variables.

Indeed, evaluating the values—different kinds of values and on different levels—of health care innovation is complex. While clinical testing can demonstrate its usefulness to end users, it is often difficult to determine whether the outcomes involve soft values (eg, reduced distress and improved health literacy and self-efficacy) or hard, monetary values [14]. Furthermore, the distribution of costs and value resulting from an innovation can be intricate, making it hard to assess. Questions arise about the initial and recurrent costs and whether they relate to the product or to new services that require additional staff tasks. Early in the innovation process, there is a need to specify both costs and values, be they monetary or qualitative. Clearly describing and anchoring a value proposition, whether it involves soft or hard values, with stakeholders early in the process is crucial for understanding costs and outcomes. However, finding effective ways to evaluate an innovation before it is ready for large-scale testing can be challenging. Similarly, value and costs stemming from an innovation can be distributed across the organization or organizations in ways that are difficult to assess. Calculating the health costs of improving care processes that involve many actors in a complex organization such as the VGR is complicated [61]. More pragmatic evaluations of both effectiveness and process outcomes are needed and can help show the effect from different angles [53].

For all cases dealt with in this study, the value proposition in terms of quality of life and cost reduction needs further evaluation over a longer period relying on data related to care costs for at-risk patient groups. The D-Foot database contains valuable information on risk groups and foot status, serving as a data source for audits and evaluations to optimize foot care. It could also function as a quality registry, a new diabetic foot register in Sweden. In the MoodMapper, the users comprise patients; their clinical teams; and, occasionally, relatives or caregivers. A published study highlights the value of implementing and receiving psychological relapse prevention for these groups, leading to improved understanding of bipolar disorder [62] that might, in turn, lead to enhanced working relationships and better condition management. However, the evidence is not consistent, and further studies are needed [61]. Moreover, for patients with bipolar disorders, having some of their behavioral patterns (such as step count and estimated sleep) automatically monitored meant that there needed to be a great deal of trust in how data are handled, something that might be difficult for patients experiencing symptoms of paranoia.

Co-Design and Coproduction

Involving users both directly and indirectly at an early stage of the development process is highly beneficial, particularly because what benefits one person may pose challenges for another, thereby creating complexity. Although there are several examples of how coproduction is useful in the innovation process, the existence of complexity must not be neglected in the co-design.

Bottom-up innovations in care encompass a wide spectrum of patient-centric approaches, empowering individuals and communities to actively participate in projects aiming to support well-being. These innovations, driven by the challenges that health care faces, range from self-management tools [62-64] and patient support networks to community-driven health programs [6-8,10,59-61,65-67]. They appear with different approaches, such as lean production [9] and Six Sigma [68]. Coproduction can enhance the 3 Rs in research—reach, rigor, and relevance [69]—by ensuring that the right needs are addressed and that the innovation is practical for both patients and staff.

Enthusiastic innovators and staff should be engaged early in the process, along with representatives from patient organizations or individuals with relevant experience. There is a strong movement toward involving patients in health care improvement, and genuine engagement is necessary for truly bottom-up innovation involvement [70] as it can lead to more radical solutions or suggestions when used correctly [71]. If a technological innovation is too demanding or unfamiliar for users, it is unlikely to be accepted. Piloting with stakeholders is crucial for assessing practicality [59], and using input from stakeholders in the right phase can increase the possibility of finding radical suggestions, as well as saving time for both parties (developers and patients) [71]. We support the idea that coproduction incorporating the multifaceted aspects of complexity is necessary in the evaluation of success in the implementation of bottom-up innovations [4].

XSL•FO
RenderX

Methodological Considerations

Performing a retrospective, deductive analysis as a case study [21] with 4 cases with differences regarding where they were in the innovation process and with different technical solutions was challenging, but it provided multiple valuable insights. The authors found that, before using the NASSS-CAT, users need to be familiar with the NASSS framework [15]. The NASSS-CAT appeared deceptively easy at first, but it was more difficult to use and more time-consuming than expected. A need for a way to track how we could jointly understand and agree on the meaning of the NASSS-CAT by using CCA became apparent during the work, leading to a common language being agreed upon and a consensus being reached on how to interpret the terminology used in the NASSS-CAT. During the CCA, discussions about how to interpret the questions in the 7 domains of the NASSS-CAT took place in cycles, and thus, it was a continuous learning process. As intended by the method [46,47], finding disagreements and negotiating led to a higher degree of understanding not just of the instrument but also of the concept of complexity. The 4 innovators contributed multiple perspectives based on their own cases and discussed their different understandings of the narratives, the domain questions, and the subquestions. As the authors used a nontranslated version of the NASSS-CAT and are native speakers of Swedish and not English, the CCA helped them understand the questions in the NASSS-CAT. Therefore, the use of CCA statements and negotiations on how to interpret the questions in the NASSS-CAT facilitated the analysis and helped create a common language within the group.

The NASSS framework was developed through a detailed review of the existing literature and clinical cases [15,20], but to our knowledge, the tools (NASSS-CAT) have so far been sparsely tested for their ability to unveil complexity in bottom-up projects in public health care. Going from commonly used methods for quality improvement (eg, using the Plan-Do-Study-Act method [72] to incorporate complexity assessment) shows promising results. A recent study used the framework and tool combined with the Plan-Do-Study-Act cycles of improvement to plan and evaluate digital services for patients in Sweden [73]. Similar to our retrospective analysis, that study identified several elements of complexity, explaining a gap among the capacity of adopters, the organization, the wider system, and how intended users valued the service. This gap hindered the innovations from integrating new services into routine care effectively [73]. Similarly to us, these authors found the tool and framework helpful in that they allowed for deeper insights into the project compared to only following method, approach, or cycles or other tools or models for innovation. It seems that, even if complexity is revealed early in the process, this still does not solve the problems. However, if people working with innovation or in supporting innovation become more aware of complex elements, issues might be easier to anticipate or even deal with earlier. Such awareness can thereby help explain obstacles and prevent failure, hence enabling more successful innovation projects in health care, as presented by Greenhalgh et al [15].

Strengths and Limitations

The strength of this study lies in the 4 different cases representing both somatic and psychiatric care and the innovators' long experience in both health care and eHealth. The diversity of innovations presented and the different departments that each of the innovators worked in contribute to a broad overview of shared experiences. None of the innovators had worked together before this study. The fact that all cases came from the same region with the same support function strengthens our results by showing that (1) knowledge of complexity needs to be improved in such systems and (2) the project itself contains complexity in different domains even if we found several common problems. Therefore, the study increased the understanding of the role of complexity, not only in the studied bottom-up innovations but also in the system in which the innovations took place, through prolonged engagement [50]. This study was strengthened by the support of one of the authors, who was involved in the development of the NASSS-CAT [20], but despite this, it appears that adaptation to the setting (geographic and cultural) is crucial.

The retrospective NASSS-CAT analysis of 3 of the 4 cases was mainly performed by the respective innovators without direct input from stakeholders involved in each of the cases. This meant that only 1 perspective from the many actors involved in each of the projects was put forward. The rationale for this was that each innovator had already faced and, therefore, was acquainted with the diverse complexities addressed in all 7 domains. However, other perspectives might have further improved the analysis. The PoC Dashboard project was assessed regarding complexity in a workshop with stakeholders and discussed with management [54].

Even though the NASSS-CAT tools have been used previously [74], more testing in clinical bottom-up innovation cases is needed to scrutinize their utility in a Swedish setting and to learn from the experiences originating from 4 different cases that used the tools.

Use and Usefulness of the NASSS-CAT

In this section, experiences of the use and utility of the NASSS-CAT are presented. At the start, the 4 bottom-up innovators were naïve and expected the NASSS-CAT [20] to be easy to comprehend and use as they identified complexities in their own innovations. As mentioned previously, by using the CCA interpretations from multiple perspectives (the 4 cases), a shared understanding and language regarding how to interpret the questions in the NASSS-CAT was established. The results from the CCA revealed that each of the 4 innovators needed to clarify or consider a number of points in their own NASSS-CAT analysis while assessing complexities in each of the domains. The most important issues were as follows:

1. To define the time frame and the scope that the innovator is assessing.
2. To define the intended users and adopters at the time of the studied project.
3. To rethink the way in which the value proposition can be measured.

XSL•FO
RenderX

4. To consider that questions in the NASSS-CAT regarding ownership; supply chains; and use and spread at the local, regional, national, and international level belong to both "Domain 5: organization" and "Domain 2: technology."

Moving Forward in Supporting Bottom-Up Innovation

This study explored insights from the NASSS framework, revealing that the adoption and dissemination of technological innovations are influenced by organizational and systemic factors rather than by the innovations themselves. The success of bottom-up innovators in navigating complexities emphasizes common challenges across innovations. The NASSS framework has illuminated various perspectives that can either facilitate or impede the adoption, scale-up, and dissemination of technological innovations.

The 4 bottom-up innovators managed to navigate through the complexities within the innovative system, uncovering overarching challenges that unified their respective innovations. However, it is essential to recognize that the NASSS-CAT cannot be used as a linear checklist. Existing support systems, while aiming to foster innovation, may unintentionally follow a linear approach rather than embracing frameworks suitable for complex interventions, such as the Medical Research Council guidance [75]. To better support health care innovators, a "midway filter system" is needed, which offers profound insights into innovation within complex systems. Implementing such a filter between top-down and bottom-up approaches would facilitate bidirectional knowledge transfer. It would enable clinical insights, ideas, and innovations to be discussed in harmony with the regulatory framework, ultimately leading to improved and equitable health care as envisioned by Tierney et al [10], who found that localized, regional, and flexible innovations can shape care in the future [10]. Incentives to connect bottom-up initiatives with a top-down vision at a national level in building systems for digital innovation and health IT are presented by Sheik et al [67] from the United Kingdom. We share their vision to improve usability and interoperability and integrate bottom-up with top-down resources.

Our study, similarly to the research by Batalden and Davidoff [76], discusses the complexities of integrating grassroots idea innovations into established health care systems. Batalden and Davidoff [76] highlight the need for organizational changes and a shift in culture to recognize the value of patient-driven innovations and effectively incorporate them into clinical practice.

Future studies should consider a translation project of the NASSS framework from English into Swedish. This would facilitate the framework's use in a Swedish context, similar to the translations of health-related quality of life questionnaires, which follow guidelines to ensure validity in terms of language and culture [53]. In addition, an evaluation is recommended alongside updates to the NASSS-CAT. Some subquestions may benefit from further splitting, such as assessing the likelihood of technology obsolescence or the measurement of alternative ways to evaluate innovation. It is crucial to involve relevant stakeholders in these changes. Cultural adaptation should receive significant emphasis to provide a language that is relevant to the Swedish context. We also suggest that future studies explore similarities and differences regarding the existence of complexities when bottom-up innovations are developed and implemented in other regions. Finally, we consider making a follow-up prospective evaluation of our 4 innovations. By doing this, we can possibly review the impact of this study on the long-term outcomes of each innovation using the NASSS-CAT LONG.

The Progress of the 4 Cases

The insights gained from our exploration of the existence of complexities in innovation processes led to some of the presented innovations being appreciated in the VGR. The Digi-Do and the D-Foot have gradually, during the study, been acknowledged as important in building future care with digital tools. The VGR has granted the innovator of the D-Foot the legal rights to be spread nationally and internationally and to be implemented and scaled up to prevent DFUs through early screening.

The Digi-Do has been evaluated, and the results show a positive effect on the users, indicating reduced levels of distress and an improved sense of preparedness [43,77]. Hence, the difficulties in evaluating soft values have been successfully dealt with. Since the analysis presented in this paper, the intellectual properties have been transferred to the VGR together with the RT department, and updated versions of the Digi-Do are underway.

Learnings from the complexity assessment of the PoC Dashboard [54] helped address the challenges differently by going for a simpler technical solution with less dependencies on other information systems and focusing on core features such as supporting patients and health care professionals in the planning and evaluation of care. This was done by adapting Dialog+, which is both a tool to measure and monitor patient-reported outcome measures and patient-reported experience measures and a solutions-focused methodology, to fit Swedish psychiatric care [78]. It has since then been piloted and tested for >4 different patient groups in mental health care settings and is being implemented as part of routine psychosis care. The MoodMapper is not an active innovation project in Sweden. However, it is used internationally in research to map behavior changes in mental disorders [79,80].

Conclusions

The NASSS framework increased the bottom-up innovators' understanding of the role of complexity in their innovations. The analysis provided valuable insights by identifying and bringing attention to complexities, particularly within the broader system, albeit requiring a deep understanding. This study enriched our comprehension of the pervasive role of complexity in bottom-up innovations within public health care and shed light on the practical utility of the NASSS-CAT. Early use of a validated tool aids in identifying complexities and pinpointing the domains in which these complexities exist. Importantly, even when bottom-up innovations arise within the same support organization, the complexity can vary based on the developmental phase and the unique characteristics of each project. Identifying, defining, and understanding complexity

XSL•FO
RenderX

may not solve the issues but substantially improves the prospects for successful innovation implementation provided the right expertise is available to support the process.

Successful innovation within complex organizational structures necessitates a comprehensive understanding and an adaptive leadership to surmount cultural resistance and organizational impediments. A rigid, linear, and stepwise approach risks disregarding interconnected variables and dependencies, leading to suboptimal outcomes. Success lies in embracing the complexity with its uncertainty, nurturing creativity, and adopting a nonlinear methodology that accommodates the iterative nature of innovation processes within complex organizations.

Acknowledgments

The authors are most grateful for the collaboration with the Innovation Platform in Region Västra Götaland. They would also like to acknowledge the Regional Cancer Centre West, the Department of Prosthetics and Orthotics at Sahlgrenska University Hospital, and the Department of Schizophrenia Spectrum Disorders at Sahlgrenska University Hospital for their support.

Data Availability

The data sets generated during and analyzed during this study are available from the corresponding author on reasonable request.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Description of each of the 4 cases of bottom-up innovations in Swedish health care, including a presentation of how the innovations relate to the 7 domains of the nonadoption, abandonment, scale-up, spread, and sustainability complexity assessment tool.
[DOCX File , 25 KB - humanfactors_v11i1e50889_app1.docx]

Multimedia Appendix 2
A constant comparative analysis of how to interpret the 7 domains of the nonadoption, abandonment, scale-up, spread, and sustainability complexity assessment tool.
[DOCX File , 22 KB - humanfactors_v11i1e50889_app2.docx]

Multimedia Appendix 3
Complexities in the 4 bottom-up innovations assessed using the nonadoption, abandonment, scale-up, spread, and sustainability complexity assessment tool.
[DOCX File , 24 KB - humanfactors_v11i1e50889_app3.docx]

References

1. Chan S. Complex adaptive systems. Massachusetts Institute of Technology. 2001. URL: https://web.mit.edu/esd.83/www/notebook/Complex%20Adaptive%20Systems.pdf [accessed 2024-04-08]
2. Greenhalgh T, Papoutsi C. Studying complexity in health services research: desperately seeking an overdue paradigm shift. BMC Med 2018 Jun 20;16(1):95 [FREE Full text] [doi: 10.1186/s12916-018-1089-4] [Medline: 29921272]
3. Tsoukas H. Don't simplify, complexify: from disjunctive to conjunctive theorizing in organization and management studies. J Manage Stud 2016 Jul 11;54(2):132-153. [doi: 10.1111/joms.12219]
4. Jacob C, Lindeque J, Klein A, Ivory C, Heuss S, Peter MK. Assessing the quality and impact of eHealth tools: systematic literature review and narrative synthesis. JMIR Hum Factors 2023 Mar 23;10:e45143 [FREE Full text] [doi: 10.2196/45143] [Medline: 36843321]
5. Eldh AC, Sverker A, Bendtsen P, Nilsson E. Health care professionals' experience of a digital tool for patient exchange, anamnesis, and triage in primary care: qualitative study. JMIR Hum Factors 2020 Dec 14;7(4):e21698 [FREE Full text] [doi: 10.2196/21698] [Medline: 33315014]
6. Klundert JV, de Korne D, Yuan S, Wang F, van Wijngaarden J. 'Hybrid' top down bottom up health system innovation in rural China: a qualitative analysis. PLoS One 2020;15(10):e0239307 [FREE Full text] [doi: 10.1371/journal.pone.0239307] [Medline: 33027287]
7. Lambert AS, Ces S, Malembaka EB, Van Durme T, Declercq A, Macq J. Evaluation of bottom-up interventions targeting community-dwelling frail older people in Belgium: methodological challenges and lessons for future comparative effectiveness studies. BMC Health Serv Res 2019 Jun 24;19(1):416 [FREE Full text] [doi: 10.1186/s12913-019-4240-9] [Medline: 31234857]

XSL•FO
RenderX

8. Orgill M, Gilson L, Chitha W, Michel J, Erasmus E, Marchal B, et al. A qualitative study of the dissemination and diffusion of innovations: bottom up experiences of senior managers in three health districts in South Africa. Int J Equity Health 2019 Mar 29;18(1):53 [FREE Full text] [doi: 10.1186/s12939-019-0952-z] [Medline: 30925878]
9. van Rossum L, Aij KH, Simons FE, van der Eng N, Ten Have WD. Lean healthcare from a change management perspective. J Health Organ Manag 2016 May 16;30(3):475-493 [FREE Full text] [doi: 10.1108/JHOM-06-2014-0090] [Medline: 27119398]
10. Tierney E, Hannigan A, Kinneen L, May C, O'Sullivan M, King R, et al. Interdisciplinary team working in the Irish primary healthcare system: analysis of 'invisible' bottom up innovations using Normalisation Process Theory. Health Policy 2019 Nov;123(11):1083-1092 [FREE Full text] [doi: 10.1016/j.healthpol.2019.09.002] [Medline: 31575445]
11. Greenhalgh T, Robert G, Macfarlane F, Bate P, Kyriakidou O. Diffusion of innovations in service organizations: systematic review and recommendations. Milbank Q 2004;82(4):581-629 [FREE Full text] [doi: 10.1111/j.0887-378X.2004.00325.x] [Medline: 15595944]
12. Van de Ven A, Polley DE, Garud R, Venkataraman S. The Innovation Journey. Oxford, UK: Oxford University Press; Feb 21, 2008.
13. Braithwaite J, Glasziou P, Westbrook J. The three numbers you need to know about healthcare: the 60-30-10 challenge. BMC Med 2020 May 04;18(1):102 [FREE Full text] [doi: 10.1186/s12916-020-01563-4] [Medline: 32362273]
14. Sturmberg JP, Bircher J. Better and fulfilling healthcare at lower costs: the need to manage health systems as complex adaptive systems. F1000Res 2019 Jun 5;8:789 [FREE Full text] [doi: 10.12688/f1000research.19414.1] [Medline: 31839925]
15. Greenhalgh T, Wherton J, Papoutsi C, Lynch J, Hughes G, A'Court C, et al. Beyond adoption: a new framework for theorizing and evaluating nonadoption, abandonment, and challenges to the scale-up, spread, and sustainability of health and care technologies. J Med Internet Res 2017 Nov 01;19(11):e367 [FREE Full text] [doi: 10.2196/jmir.8775] [Medline: 29092808]
16. Cohn S, Clinch M, Bunn C, Stronge P. Entangled complexity: why complex interventions are just not complicated enough. J Health Serv Res Policy 2013 Jan 01;18(1):40-43. [doi: 10.1258/jhsrp.2012.012036] [Medline: 23393041]
17. Greenhalgh T, Wherton J, Papoutsi C, Lynch J, Hughes G, A'Court C, et al. Analysing the role of complexity in explaining the fortunes of technology programmes: empirical application of the NASSS framework. BMC Med 2018 May 14;16(1):66 [FREE Full text] [doi: 10.1186/s12916-018-1050-6] [Medline: 29754584]
18. Rogers EM. Diffusion of innovations: modifications of a model for telecommunications. In: Stoetzer MW, Mahler A, editors. Die Diffusion von Innovationen in der Telekommunikation. Berlin, Heidelberg: Springer; 1995.
19. Robert G, Greenhalgh T, MacFarlane F, Peacock R. Adopting and assimilating new non-pharmaceutical technologies into health care: a systematic review. J Health Serv Res Policy 2010 Oct;15(4):243-250. [doi: 10.1258/jhsrp.2010.009137] [Medline: 20592046]
20. Greenhalgh T, Maylor H, Shaw S, Wherton J, Papoutsi C, Betton V, et al. The NASSS-CAT tools for understanding, guiding, monitoring, and researching technology implementation projects in health and social care: protocol for an evaluation study in real-world settings. JMIR Res Protoc 2020 May 13;9(5):e16861 [FREE Full text] [doi: 10.2196/16861] [Medline: 32401224]
21. Yin RK. Choices in designing qualitative research studies. In: Yin RK, editor. Qualitative Research from Start to Finish. New York, NY: The Guilford Press; 2003:359-386.
22. Västra Götalandsregionen homepage. Västra Götalandsregionen. URL: https://www.vgregion.se/ [accessed 2024-01-29]
23. Armstrong DG, Tan TW, Boulton AJ, Bus SA. Diabetic foot ulcers: a review. JAMA 2023 Jul 03;330(1):62-75. [doi: 10.1001/jama.2023.10578] [Medline: 37395769]
24. Ibrahim A, Jude E, Langton K, Martinez-De Jesus FR, Harkless LB, Gawish H, et al. IDF clinical practice recommendations on the diabetic foot – 2017. International Diabetes Federation. 2017. URL: https://idf.org/media/uploads/2023/05/attachments-61.pdf [accessed 2024-04-08]
25. Fotundersökning vid diabetes. Sveriges Kommuner och Landsting. 2018. URL: https://swenurse.se/download/18.32612fa517b3be241ddb892d/1629654183409/fotundersokning-vid-diabetes.pdf [accessed 2024-04-08]
26. Personcentrerat och sammanhållet vårdförlopp diabetes med hög risk för fotsår. Nationellt Programområde Endokrina Sjukdomar. 2022 Jun 3. URL: https://d2flujgsl7escs.cloudfront.net/external/Vardforlopp_Diabetes.pdf [accessed 2024-04-08]
27. Hellstrand Tang U, Zügner R, Lisovskaja V, Karlsson J, Hagberg K, Tranberg R, et al. The D-Foot, for prosthetists and orthotists, a new eHealth tool useful in useful in risk classification and foot assessment in diabetes. Foot Ankle Online J 2017 Aug;10(2):4 [FREE Full text] [doi: 10.3827/faoj.2017.1002.0004]
28. Zügner R, Jarl G, Sundberg L, Tang UH. Experiences of using a digital tool, the D-foot, in the screening of risk factors for diabetic foot ulcers. J Foot Ankle Res 2022 Dec 13;15(1):90 [FREE Full text] [doi: 10.1186/s13047-022-00594-9] [Medline: 36514099]
29. Ribu L, Hanestad BR, Moum T, Birkeland K, Rustoen T. A comparison of the health-related quality of life in patients with diabetic foot ulcers, with a diabetes group and a nondiabetes group from the general population. Qual Life Res 2007 Mar;16(2):179-189. [doi: 10.1007/s11136-006-0031-y] [Medline: 17091370]
30. Regulation (EU) 2017/745 of the European parliament and of the council. Official Journal of the European Union. 2017 Apr 5. URL: https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A32017R0745 [accessed 2024-04-08]

XSL•FO
RenderX

31. Ragnarson Tennvall G, Apelqvist J. Health-economic consequences of diabetic foot lesions. Clin Infect Dis 2004 Aug 01;39 Suppl 2:S132-S139. [doi: 10.1086/383275] [Medline: 15306992]
32. Hellstrand S, Sundberg L, Karlsson J, Zügner R, Tranberg R, Hellstrand Tang U. The value of time consumed in healthcare – prevention of diabetic foot ulcers. Heliyon Preprint posted online November 1, 2023 [FREE Full text] [doi: 10.2139/ssrn.4610118]
33. Hellstrand S, Hellstrand Tang U. A new approach to quantifying the sustainability effects of healthcare: applied to the diabetic foot. Foot Ankle Online J 2019;12(3):5. [doi: 10.3827/faoj.2019.1203.0005]
34. Pontin E, Peters S, Lobban F, Rogers A, Morriss RK. Enhanced relapse prevention for bipolar disorder: a qualitative investigation of value perceived for service users and care coordinators. Implement Sci 2009 Feb 09;4(1):4 [FREE Full text] [doi: 10.1186/1748-5908-4-4] [Medline: 19203373]
35. Swedish Medical Products Agency. Manufacture of medical devices 2024. URL: https://www.lakemedelsverket.se/sv [accessed 2024-04-08]
36. Regional utvecklingsstrategi för Västra Götaland 2021 – 2030. Västra Götalandsregionen. 2021. URL: https://tinyurl.com/266rt7be [accessed 2023-12-26]
37. Millennium. Västra Götalandsregionen. URL: https://www.vgregion.se/halsa-och-vard/vardgivarwebben/vardskiftet/millennium/ [accessed 2024-01-29]
38. Manufacture of medical devices. Swedish Medical Products Agency. 2020. URL: https://www.lakemedelsverket.se/en/medical-devices/manufacture#hsecondarybody3 [accessed 2024-04-02]
39. Nationella riktlinjer för diabetesvård. Socialstyrelsen. URL: https://www.socialstyrelsen.se/kunskapsstod-och-regler/regler-och-riktlinjer/nationella-riktlinjer/riktlinjer-och-utvarderingar/diabetes/ [accessed 2024-04-08]
40. Schaper NC, van Netten JJ, Apelqvist J, Bus SA, Fitridge R, Game F, et al. Practical guidelines on the prevention and management of diabetes-related foot disease (IWGDF 2023 update). Diabetes Metab Res Rev 2024 Mar;40(3):e3657. [doi: 10.1002/dmrr.3657] [Medline: 37243927]
41. Ragnarson Tennvall G, Apelqvist J. Health-related quality of life in patients with diabetes mellitus and foot ulcers. J Diabetes Complications 2000;14(5):235-241. [doi: 10.1016/s1056-8727(00)00133-1] [Medline: 11113684]
42. Ragnarson Tennvall G, Apelqvist J. Prevention of diabetes-related foot ulcers and amputations: a cost-utility analysis based on Markov model simulations. Diabetologia 2001 Nov;44(11):2077-2087. [doi: 10.1007/s001250100013] [Medline: 11719840]
43. Grynne A, Browall M, Fristedt S, Ahlberg K, Smith F. Integrating perspectives of patients, healthcare professionals, system developers and academics in the co-design of a digital information tool. PLoS One 2021 Jul 16;16(7):e0253448 [FREE Full text] [doi: 10.1371/journal.pone.0253448] [Medline: 34270564]
44. Fristedt S, Smith F, Grynne A, Browall M. Digi-Do: a digital information tool to support patients with breast cancer before, during, and after start of radiotherapy treatment: an RCT study protocol. BMC Med Inform Decis Mak 2021 Feb 25;21(1):76 [FREE Full text] [doi: 10.1186/s12911-021-01448-3] [Medline: 33632215]
45. Gremyr A, Andersson Gäre B, Greenhalgh T, Malm U, Thor J, Andersson AC. Using complexity assessment to inform the development and deployment of a digital dashboard for schizophrenia care: case study. J Med Internet Res 2020 Apr 23;22(4):e15521 [FREE Full text] [doi: 10.2196/15521] [Medline: 32324143]
46. Steen M. Virtues in participatory design: cooperation, curiosity, creativity, empowerment and reflexivity. Sci Eng Ethics 2013 Sep 18;19(3):945-962. [doi: 10.1007/s11948-012-9380-9] [Medline: 22806218]
47. Smith F, Wallengren C, Öhlén J. Participatory design in education materials in a health care context. Action Res 2016 May 03;15(3):310-336. [doi: 10.1177/1476750316646832]
48. Innovation i komplexa organisationer. Västra Götalandsregionen. 2021 Feb 18. URL: https://www.vgregion.se/siteassets/ovriga_webbplatser/innovationsplattformen/innovation-i-komplexa-organisationer.pdf [accessed 2024-04-08]
49. Nationellt vårdprogram bröstcancer. Kunskapsbank. URL: https://kunskapsbanken.cancercentrum.se/diagnoser/brostcancer/vardprogram/ [accessed 2024-04-08]
50. International guidelines on breast cancer care. European Commission. URL: https://healthcare-quality.jrc.ec.europa.eu/ecibc/international-guidelines [accessed 2024-04-08]
51. Branca Vergano L, Monesi M, Vicenti G, Bizzoca D, Solarino G, Moretti B. Posterior approaches in malleolar fracture: when, why and how. J Biol Regul Homeost Agents 2020;34(3 Suppl. 2):89-95. ADVANCES IN MUSCULOSKELETAL DISEASES AND INFECTIONS. [Medline: 32856446]
52. Nationella medicinska informationssystem - NMI. Läkemedelsverket. 2019 Oct 30. URL: https://www.lakemedelsverket.se/sv/medicinteknik/tillverka/nationella-medicinska-informationssystem#hmainbody1 [accessed 2023-12-26]
53. Ventura F, Brovall M, Smith F. Beyond effectiveness evaluation: contributing to the discussion on complexity of digital health interventions with examples from cancer care. Front Public Health 2022 Jul 22;10:883315 [FREE Full text] [doi: 10.3389/fpubh.2022.883315] [Medline: 35968430]
54. Gremyr A, Holmberg C, Thor J, Malm U, Gäre BA, Andersson AC. How a point-of-care dashboard facilitates co-production of health care and health for and with individuals with psychotic disorders: a mixed-methods case study. BMC Health Serv Res 2022 Dec 30;22(1):1599 [FREE Full text] [doi: 10.1186/s12913-022-08992-2] [Medline: 36585696]
55. Vision e-hälsa 2025. E-hälsa. URL: https://ehalsa2025.se/ [accessed 2024-04-08]

XSL•FO
RenderX

56. Lag (2007:1091) om offentlig upphandling. Sveriges Riksdag. URL: https://www.riksdagen.se/sv/dokument-och-lagar/dokument/svensk-forfattningssamling/lag-20071091-om-offentlig-upphandling_sfs-2007-1091/ [accessed 2024-04-08]
57. Bell H, Rees D, Huxtable-Thomas L, Rich N, Miller E, Thomas R. Innovation adoption research in healthcare: understanding context and embracing complexity. In: Proceedings of the 18th European Conference on Innovation and Entrepreneurship. 2023 Presented at: ECIE 2023; September 21-22, 2023; Porto, Portugal. [doi: 10.34190/ecie.18.1.1865]
58. Chaves BG, Briand C, Bouabida K. Innovation in healthcare organizations: concepts and challenges to consider. Int J Health Res Innov 2021;9(1):1-14. [doi: 10.47260/ijhri/911]
59. Dugstad J, Eide T, Nilsen ER, Eide H. Towards successful digital transformation through co-creation: a longitudinal study of a four-year implementation of digital monitoring technology in residential care for persons with dementia. BMC Health Serv Res 2019 Jun 10;19(1):366 [FREE Full text] [doi: 10.1186/s12913-019-4191-1] [Medline: 31182093]
60. Schliemann T, Danielsen C, Virtanen T, Vuokko R, Hardardottir GA, Alsaker MA, et al. eHealth standardisation in the Nordic countries: technical and partially semantics standardisation as a strategic means for realising national policies in eHealth. Nordic Council of Ministers. 2019. URL: https://norden.diva-portal.org/smash/get/diva2:1340369/FULLTEXT01.pdf [accessed 2024-04-08]
61. Steenkamer B, Drewes H, Putters K, van Oers H, Baan C. Reorganizing and integrating public health, health care, social care and wider public services: a theory-based framework for collaborative adaptive health networks to achieve the triple aim. J Health Serv Res Policy 2020 Jul;25(3):187-201. [doi: 10.1177/1355819620907359] [Medline: 32178546]
62. Mondéjar AG, Silva-Calpa GF, Raposo AB, Mograbi DC. BraPolar: an m-health application for remote monitoring of people with bipolar disorder. In: Proceedings of the IEEE 7th International Conference on Serious Games and Applications for Health (SeGAH). 2019 Presented at: SeGAH 2019; August 5-7, 2019; Kyoto, Japan. [doi: 10.1109/segah.2019.8882469]
63. Faurholt-Jepsen M, Kessing LV. Monitoring and treatment in patients with bipolar disorder using smartphones-new perspectives for improved quality in patient care. Psychiatry Res 2022 Nov;317:114844. [doi: 10.1016/j.psychres.2022.114844] [Medline: 36115167]
64. Shan R, Sarkar S, Martin SS. Digital health technology and mobile devices for the management of diabetes mellitus: state of the art. Diabetologia 2019 Jun;62(6):877-887. [doi: 10.1007/s00125-019-4864-7] [Medline: 30963188]
65. Coulter A, Locock L, Ziebland S, Calabrese J. Collecting data on patient experience is not enough: they must be used to improve care. BMJ 2014 Mar 26;348:g2225. [doi: 10.1136/bmj.g2225] [Medline: 24671966]
66. Coulter A, Oldham J. Person-centred care: what is it and how do we get there? Future Hosp J 2016 Jun;3(2):114-116 [FREE Full text] [doi: 10.7861/futurehosp.3-2-114] [Medline: 31098200]
67. Sheikh A, Anderson M, Albala S, Casadei B, Franklin BD, Richards M, et al. Health information technology and digital innovation for national learning health and care systems. Lancet Digit Health 2021 Jun;3(6):e383-e396 [FREE Full text] [doi: 10.1016/S2589-7500(21)00005-4] [Medline: 33967002]
68. Niñerola A, Sánchez-Rebull MV, Hernández-Lara AB. Quality improvement in healthcare: Six Sigma systematic review. Health Policy 2020 Apr;124(4):438-445. [doi: 10.1016/j.healthpol.2020.01.002] [Medline: 32156468]
69. Balazs CL, Morello-Frosch R. The three R's: how community based participatory research strengthens the rigor, relevance and reach of science. Environ Justice 2013 Feb;6(1):9-16 [FREE Full text] [doi: 10.1089/env.2012.0017] [Medline: 24260590]
70. Ocloo J, Matthews R. From tokenism to empowerment: progressing patient and public involvement in healthcare improvement. BMJ Qual Saf 2016 Aug 18;25(8):626-632 [FREE Full text] [doi: 10.1136/bmjqs-2015-004839] [Medline: 26993640]
71. Gremyr I, Elg M, Smith F, Gustavsson S. Exploring the phase for highest impact on radicality: a cross-sectional study of patient involvement in quality improvement in Swedish healthcare. BMJ Open 2018 Nov 08;8(11):e021958 [FREE Full text] [doi: 10.1136/bmjopen-2018-021958] [Medline: 30413500]
72. Taylor MJ, McNicholas C, Nicolay C, Darzi A, Bell D, Reed JE. Systematic review of the application of the plan-do-study-act method to improve quality in healthcare. BMJ Qual Saf 2014 Apr 11;23(4):290-298 [FREE Full text] [doi: 10.1136/bmjqs-2013-001862] [Medline: 24025320]
73. Sacchi C, Andersson K, Roczniewska M, Luckhaus JL, Malmqvist M, Rodmalm LP, et al. Mind the gap: analysis of two pilot projects of a home telehealth service for persons with complex conditions in a Swedish hospital. BMC Health Serv Res 2023 May 09;23(1):463 [FREE Full text] [doi: 10.1186/s12913-023-09409-4] [Medline: 37161458]
74. Greenhalgh T. How to improve success of technology projects in health and social care. Public Health Res Pract 2018 Sep 27;28(3):2831815 [FREE Full text] [doi: 10.17061/phrp2831815] [Medline: 30406256]
75. Skivington K, Matthews L, Simpson SA, Craig P, Baird J, Blazeby JM, et al. A new framework for developing and evaluating complex interventions: update of Medical Research Council guidance. BMJ 2021 Sep 30;374:n2061 [FREE Full text] [doi: 10.1136/bmj.n2061] [Medline: 34593508]
76. Batalden PB, Davidoff F. What is "quality improvement" and how can it transform healthcare? Qual Saf Health Care 2007 Feb;16(1):2-3 [FREE Full text] [doi: 10.1136/qshc.2006.022046] [Medline: 17301192]
77. Grynne A, Wångdahl J, Fristedt S, Smith F, Browall M. Women's experience of the health information process involving a digital information tool before commencing radiation therapy for breast cancer: a deductive interview study. BMC Health Serv Res 2023 Aug 09;23(1):842 [FREE Full text] [doi: 10.1186/s12913-023-09837-2] [Medline: 37559113]

XSL•FO
RenderX

78. Lindenfalk B, Gremyr A, Lundmark M, Jacobsson T. Digitally mediated schizophrenia care - a Swedish case of translating, designing and expert evaluation of dialog. Stud Health Technol Inform 2022 Jun 06;290:882-886. [doi: 10.3233/SHTI220206] [Medline: 35673145]
79. Moukhtarian TR, Reinhard I, Morillas-Romero A, Ryckaert C, Mowlem F, Bozhilova N, et al. Wandering minds in attention-deficit/hyperactivity disorder and borderline personality disorder. Eur Neuropsychopharmacol 2020 Sep;38:98-109. [doi: 10.1016/j.euroneuro.2020.07.005] [Medline: 32703662]
80. Moukhtarian TR, Reinhard I, Moran P, Ryckaert C, Skirrow C, Ebner-Priemer U, et al. Comparable emotional dynamics in women with ADHD and borderline personality disorder. Borderline Personal Disord Emot Dysregul 2021 Feb 12;8(1):6 [FREE Full text] [doi: 10.1186/s40479-021-00144-y] [Medline: 33579385]

Abbreviations

CCA: constant comparative analysis
CE: Conformité Européenne
DFU: diabetic foot ulcer
IPF: Innovation Platform
NASSS: nonadoption, abandonment, scale-up, spread, and sustainability
NASSS-CAT: nonadoption, abandonment, scale-up, spread, and sustainability complexity assessment tool
NMI: national medical information system
RT: radiation therapy
VGR: Region Västra Götaland

Edited by A Kushniruk; submitted 18.07.23; peer-reviewed by M Aanestad, J Wells; comments to author 20.11.23; revised version received 24.02.24; accepted 02.03.24; published 26.04.24.

Please cite as:
Hellstrand Tang U, Smith F, Karilampi UL, Gremyr A
Exploring the Role of Complexity in Health Care Technology Bottom-Up Innovations: Multiple-Case Study Using the Nonadoption, Abandonment, Scale-Up, Spread, and Sustainability Complexity Assessment Tool
JMIR Hum Factors 2024;11:e50889
URL: https://humanfactors.jmir.org/2024/1/e50889
doi:10.2196/50889
PMID:38669076

©Ulla Hellstrand Tang, Frida Smith, Ulla Leyla Karilampi, Andreas Gremyr. Originally published in JMIR Human Factors (https://humanfactors.jmir.org), 26.04.2024. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on https://humanfactors.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Provider Adoption of mHealth in Rural Patient Care: Web-Based Survey Study

Bryan P Weichelt[1*], MBA, PhD; Rick Burke[1*], MPH; Burney Kieke[2*], MS; Matt Pilz[1*], BS; Neel Shimpi[3*], PhD

[1]National Farm Medicine Center, Marshfield Clinic Research Institute, Marshfield, WI, United States

[2]Center for Clinical Epidemiology & Population Health, Marshfield Clinic Research Institute, Marshfield, WI, United States

[3]American Dental Association, Chicago, IL, United States

[*]all authors contributed equally

Corresponding Author:
Bryan P Weichelt, MBA, PhD
National Farm Medicine Center
Marshfield Clinic Research Institute
1000 N Oak Ave
Marshfield, WI, 54449
United States
Phone: 1 715 221 7276
Email: weichelt.bryan@marshfieldresearch.org

Abstract

Background: Physicians and patient-facing caregivers have increasingly used mobile health (mHealth) technologies in the past several years, accelerating during the COVID-19 pandemic. However, barriers and feedback surrounding adoption remain relatively understudied and varied across health systems, particularly in rural areas.

Objective: This study aims to identify provider adoption, attitudes, and barriers toward mHealth in a large, multisite, rural US health care system. We investigated (1) mHealth apps that providers use for their own benefit and (2) mHealth apps that a provider uses in conjunction with a patient.

Methods: We surveyed all patient-seeing providers within the Marshfield Clinic Health System with a brief, 16-item, web-based survey assessing attitudes toward mHealth, adoption of these technologies, and perceived barriers faced by providers, their peers, and the institution. Survey results were summarized via descriptive statistics, with log-binomial regression and accompanying pairwise analyses, using Kruskal-Wallis and Jonckheere-Terpstra tests for significance, respectively. Respondents were grouped by reported clinical role and specialty.

Results: We received a 38% (n/N=916/2410) response rate, with 60.7% (n=556) of those sufficiently complete for analyses. Roughly 54.1% (n=301) of respondents reported mHealth use, primarily around decision-making and supplemental information, with use differing based on provider role and years of experience. Self-reported barriers to using mHealth included a lack of knowledge and time to study mHealth technologies. Providers also reported concerns about patients' internet access and the complexity of mHealth apps to adequately use mHealth technologies. Providers believed the health system's barriers were largely privacy, confidentiality, and legal review concerns.

Conclusions: These findings echo similar studies in other health systems, surrounding providers' lack of time and concerns over privacy and confidentiality of patient data. Providers emphasized concerns over the complexity of these technologies for their patients and concerns over patients' internet access to fully use mHealth in their delivery of care.

(JMIR Hum Factors 2024;11:e55443) doi:10.2196/55443

KEYWORDS

mHealth; clinician; physician; rural; patient; mobile; health care; adoption; attitude; attitudes; opinion; perception; perceptions; perspective; perspectives; acceptance; mobile health; app; apps; provider; providers; physicians; survey; surveys; barrier; barriers; digital health

XSL•FO
RenderX

Introduction

Increased technological and medical advancements have naturally led to the intersection of these 2 fields of study, commonly known as mobile health (mHealth) [1]. mHealth has been defined as "mobile computing, medical sensor, and communication technologies for healthcare" and has seen increasing adoption in recent years, particularly following the shift to virtual health delivery during the COVID-19 pandemic [1-3]. However, given privacy concerns, institutional hesitancy, and the wide array of programs and devices available, adoption and use of mHealth have been mixed [4].

The 2 prevailing mobile platforms include Android and iOS, which collectively comprise more than 99% of mobile use today on phones, tablets, and wearable devices [5]. Traditionally, the development of mobile apps required maintaining separate codebases and expertise per platform. Advances in web technologies, coupled with the establishment of these 2 universal operating systems, have ushered in new and dynamically evolving cross-platform solutions. Vastly expanded broadband cellular networks have simultaneously led to a surge in mobile accessibility, with 95% of the globe reaching 3G coverage as of 2022 and 80% obtaining 4G or faster speeds, including throughout rural regions [6].

To streamline development resources while maximizing user reach, cross-platform frameworks have become dominant across all sectors of mobile apps, including mHealth [7]. These libraries leverage existing technologies, often derivatives of web languages such as HTML, cascading style sheets, and JavaScript, while seamlessly interfacing with the native capabilities of each platform. Two of the most popular architectures for modern cross-platform apps include React Native (Meta Platforms) and Flutter (Google) [8]. Similarly, 2 popular hybrid solutions include Cordova and Capacitor, which can efficiently embed existing websites into native web views to achieve familiar native app behavior.

Emerging app-connected wearables—smartwatches, eyewear, earwear, and clothing—synergize with cross-platform apps to offer new ways of interacting with consumers and patients. Adaptation of wearable technologies has boomed in recent years, with a projected 1 billion circulated wearables in 2022 compared with 325 million in 2016 [9]. Android and iOS smartwatches offer core health initiatives, with sensors and apps that can automatically analyze heart health, blood oxygen, sleep cycles, and fitness. This innovation is rapidly superseding traditional life alert functionality with recent developments including fall detection, automatic location-aware emergency dialing, predictive warnings of heart arrhythmia, and reported seamless syncing of medical records [10].

Many promising new use cases of wearable tech in the mHealth industry are emerging after years of research and pilot studies. For example, after 12 years of testing, Apple has proofed a noninvasive continuous blood glucose monitor system that uses silicon photonics and optical spectroscopy, which is expected to be miniaturized into a common watch-sized wearable within 3 years for consumer use [11]. Manufacturers have also continued exploring augmented reality medical applications for eyewear and are working on adaptive lens adjustment technologies, which would dynamically adjust to one's eyesight with no prescription lens required. Other types of wearable devices are continuously being tested, including ones that can monitor saliva or tear gland fluids to detect eye or oral diseases, among other medical conditions [12].

These advances in technologies and applications have moved at incredible speeds, most often ahead of health systems' and providers' organizational abilities or individual preparedness to adopt, test, and implement for their own use or use with patients. Nonetheless, health care providers can and do leverage available advances in medical technologies for the benefit of the patient, and we would fully expect that mHealth apps and wearable technologies are no exception.

To better understand the current environment of mHealth adoption and barriers among rural providers and patients, we sought to further explore two key topics in this study: (1) apps that providers use for their own benefit and (2) apps that a provider uses in conjunction with a patient.

Khatun et al originally described a conceptual model for mhealth readiness through the lens of a health workforce in rural Bangladesh [13]. The model was later advanced and refined by Weichelt et al in 2019, furthering discussions of the interplay of rural patients, clinicians, and their organizations in mHealth adoption [14,15]. This prior research found that the organization plays a role in impacting providers' and patients' adoption of mHealth; however, we need to first gain a deeper understanding of providers' current levels of adoption and familiarity and awareness with these new technologies.

This line of research, beginning with an assessment and inventory of mHealth adoption, is essential for the future of health care delivery. Marshfield Clinic Health System (MCHS) is a predominantly rural health system with patients scattered across northern Wisconsin, the Upper Peninsula of Michigan, and beyond [16]. While well positioned to test and deploy new and innovative technologies in the broad field of mHealth, leadership first needs to gain a deeper understanding of the system's provider and patient needs, desires, and current use. Therefore, we conducted a survey of all patient-seeing providers within MCHS to identify mHealth adoption, attitudes, and perceived barriers to use.

Methods

Data Collection

In July of 2020, we emailed a survey to 2410 MCHS providers via an information systems–supplied "MCHS Providers" email list. The survey was designed to assess providers' motivators and barriers to the adoption of mHealth technologies in patient care. The survey was open and available from July 21 to August 31 (6 weeks), with 2 reminders sent every 2 weeks.

Instrument design and line of inquiry leveraged previous work by this research team, including the previously published conceptual model for assessing necessary conditions for rural health care's mHealth readiness, with an emphasis on clinician-perceived barriers. Providers were asked about

XSL•FO
RenderX

mHealth use, both personal and with patients, as well as personal, perceived colleague, patient, and institutional barriers to mHealth adoption. Providers were also asked about the perceived COVID-19 impact on mHealth use and anticipated future mHealth use after the pandemic subsided.

Incentives

Participants were presented with the option of selecting one of five local nonprofits to receive a US $10 donation for their voluntary participation in the survey. We distributed our full budgeted allotment of US $2800 as chosen by the research participants. No other incentives were offered during the study.

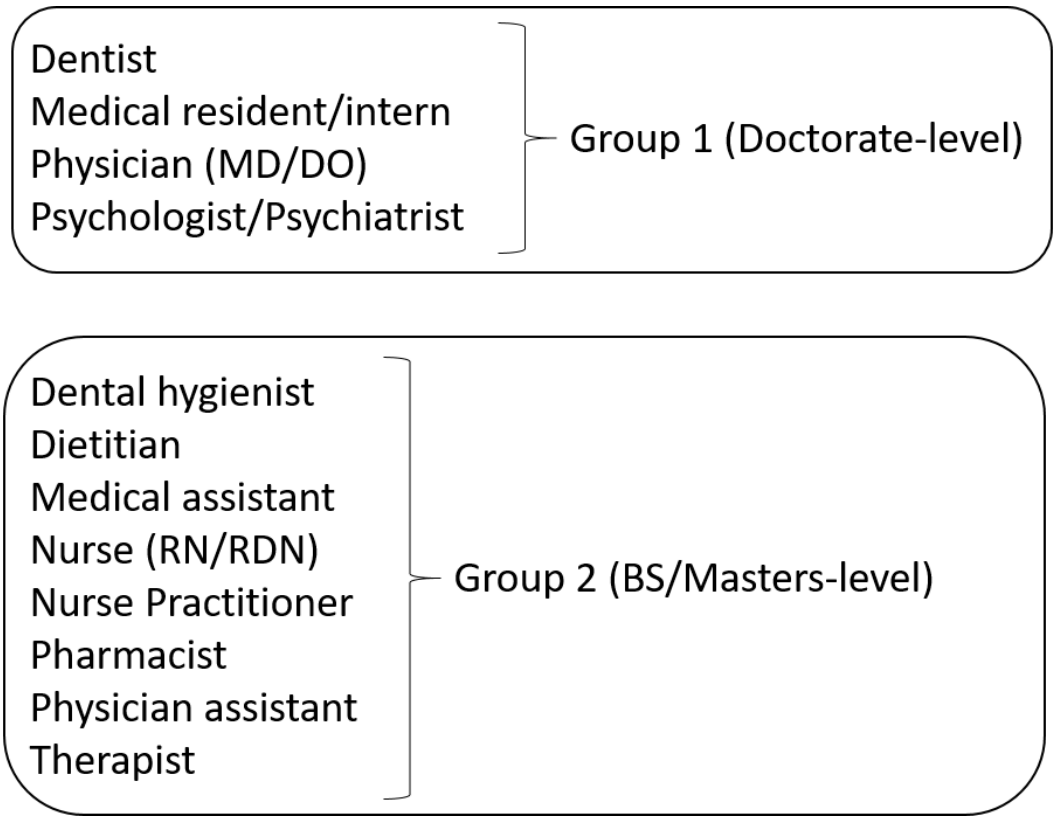
Ethical Considerations

The project submission was evaluated by the Marshfield Clinic institutional review board. It was determined that the activity as described does not meet the definition of human participant research, and no further institutional review board action was needed.

Analyses

Due to the small number of responses in some niche roles and specialties, participants' roles were grouped into 2 categories based on education and degree level (Figure 1). Provider specialty was also compartmentalized into 9 categories, mirroring the distribution of specialties across MCHS.

Figure 1. Grouping of provider roles. RDN: Registered Dietitian Nutritionist; RN: registered nurse.



We used log-binomial regression to analyze survey questions with dichotomous (yes or no) responses [17]. Specifically, we fit univariable models where the dependent variable was the dichotomous response and the independent variable was either provider role, provider specialty, or number of years in practice (7 categories: 0-5, 6-10, 11-15, 16-20, 21-25, 26-30, and >30 years). We assessed the overall statistical significance of the independent variable and proceeded with pairwise comparisons versus a referent category when warranted (ie, when the *P* value for the overall effect was ≤.05). BS and MS-level providers, family medicine, and 16-20 years in practice were the referent categories in the calculations. Since provider roles had 2 categories, corresponding pairwise comparisons were unnecessary (redundant).

A similar general strategy was used for Likert-scaled survey questions (ie, assessment of the overall statistical significance of the independent variable followed by pairwise comparisons vs a referent category when warranted). The Kruskal-Wallis test was used to evaluate overall significance and the Jonckheere-Terpstra test for pairwise comparisons [18,19]. The same independent variables were examined.

Results

Overview

We received a total of 916 responses (38% response rate), of which 556 (60.7%) responses were sufficiently complete to be included in the statistical analyses. Of these responses, 301 (54.1%) participants reported using health-related apps on their phone or tablet. The most common purposes for these apps were use as informational resources (234/301, 77.7%) and for decision-making (180/301, 59.8%). Providers who used mHealth with patients (202/556, 36.3%) reported doing so primarily for exercise and activity monitoring (105/202, 52%), and enhancing patients' experiences via the My Marshfield Clinic app (105/202, 52%), an in-house app that allows patients to schedule

XSL•FO
RenderX

appointments, view lab results, message providers, etc. Those who do not use health-related apps (255/556, 45.9%) stated their primary reasons as inadequate information available on the use of such apps (59/255, 23.1%), not having enough time to use the apps (55/255, 21.6%), or being unsure of the organization's attitudes toward mHealth (59/255, 23.1%). The most common barriers to mHealth adoption cited by providers were a lack of both knowledge about mHealth technologies (293/556, 52.7%) and time (201/556, 36.2%), as well as being unsure of their patients' access to reliable internet services (171/556, 30.8%). These same concerns arose when we asked respondents which barriers they thought other providers had surrounding mHealth (319/556, 57.4%; 283/556, 50.9%; and 163/556, 29.3%). Perceived organizational barriers to clinicians using mHealth in their practices were primarily concerns related to confidentiality (313/556, 56.3%) and mHealth technologies being too complicated for patients (288/556, 51.8%). Overall, however, providers had a favorable view of mHealth, with a majority stating they either intend to continue using mHealth following the COVID-19 pandemic or would look further into mHealth technologies.

Provider Demographics

We had a broad representation of specialties and experience levels in our responses. The most common specialties were family medicine, surgery, pediatrics, and physical and occupational therapy. The survey respondents averaged 19 years of experience practicing medicine.

Clinician Adoption

Clinician adoption of mHealth varied by role and specialty. Doctoral-level providers reported higher mHealth use on their own devices compared with other providers, with 65% (n/N=154/237) and 46% (n/N=138/300) adoption (P<.001), respectively. Among mHealth users, doctoral-level providers used these apps as an informational resource at a higher rate (131/154, 85.1% vs 98/138, 71%; P=.005). Compared with midtenure providers (16-20 years of experience, 51/81, 63% mHealth adoption), mHealth adoption levels were reported to be lower among more experienced providers (34/72, 47.2% for 21-25 years of experience; 29/64, 45.3% for 26-30 years; and 43/84, 51.2% for >30 years), and similar among less experienced providers (52/95, 54.7% for 0-5 years of experience; 43/73, 58.9% for 6-10 years; and 39/64, 60.9% for 11-15 years). mHealth use with patients was similar between doctoral-level and other providers (84/237, 35.4% vs 110/300, 36.7%; P=.77) and across the range of years of experience (33/95, 34.7% for 0-5 years of experience; 27/73, 37% for 6-10 years; 22/64, 34.4% for 11-15 years; 29/81, 35.8% for 16-20 years; 34/72, 47.2% for 21-25 years; 24/64, 37.5% for 26-30 years; and 29/84, 34.5% for >30 years; P=.63). Notably, compared with family medicine with 48.4% (n/N=31/64) mHealth use with patients, 3 specialties reported use of ≤30% (Cancer Care and Research, 6/27, 22.2%, P=.04 vs family medicine; Cardiology, 7/29, 24.1%, P=.05; Surgery, 14/60, 23.3%, P=.007), while psychiatry and psychology reported 78.3% (n/N=18/23) adoption, significantly higher than family medicine (P=.005). No important differences in reported mHealth use with patients were observed regarding diet and nutrition tracking, weight management, dental reminders, direct communication with the patient's care team, and medication reminders. Not surprisingly, psychiatry and psychology reported use of mHealth more frequently for mood and depressive symptom monitoring (12/18, 66.7% vs ≤7% for all other specialties that responded to this question [no responses in Cancer Care and Research, Cardiology, OB/GYN, Physical and Occupational Therapy, and Surgery; 1/24, 4.2% in Pediatrics; and 4/65, 6.2% in other specialties], P=.001 vs family medicine, 2/31, 6.4%) and sleep tracking (7/18, 38.9% vs ≤16.7% for all other specialties that responded to this question [no responses in Cardiology, OB/GYN and Physical and Occupational Therapy; 1/6, 16.7% in Cancer Care and Research; 2/24, 8.3% in Pediatrics; 1/14, 7.1% in Surgery; and 4/65, 6.2% in other specialties], P=.02 vs family medicine, 2/31, 6.4%). Physical and occupational therapy, cardiology, and psychiatry and psychology reported substantially higher mHealth use with patients for informational and educational purposes (12/17, 70.6%; 4/7, 57.1%; and 10/18, 55.6%, with P=.002, .048, and .02 vs family medicine, 7/31, 22.6%).

Clinicians' Perceived Barriers Category 1—Personal (Clinician)

Overall, providers reported lack of knowledge about mHealth technologies (293/556, 52.7% for themselves; 319/556, 57.4% in their perceptions regarding other clinicians) and lack of time (201/556, 36.2% and 283/556, 50.9%) as the primary personal barriers. Insufficient levels of patient internet access were also a commonly cited concern (171/556, 30.8% and 163/556, 29.3%). We found relatively few differences between provider roles and specialties regarding personal barriers to mHealth adoption. Doctoral-level providers cited a greater number of financial barriers surrounding a lack of value in mHealth technologies (29/237, 12.2% vs 12/300, 4%, P<.001 for themselves; 44/237, 18.6% vs 26/300, 8.7%, P=.001 in their perceptions regarding other clinicians), insufficient reimbursement options (27/237, 11.4% vs 14/300, 4.7%; P=.005 for themselves), and mHealth technologies not being worth the cost of adoption (22/237, 9.3% vs 7/300, 2.3%; P=.001 for themselves). With respect to their perceptions regarding other clinicians, cancer care and research providers reported a lack of communication between providers at a substantially higher rate than all other specialties (12/27, 44.4% vs 11.7%-33.3% [5/29, 17.2% in Cardiology; 9/27, 33.3% in OB/GYN; 9/47, 19.1% in Pediatrics; 6/42, 14.3% in Physical and Occupational Therapy; 5/23, 21.7% in Psychiatry and Psychology; 7/60, 11.7% in Surgery; and 45/214, 21% in other specialties] P=.02 vs family medicine [13/64, 20.3%]). Furthermore, regarding their perceptions of other clinicians, OB and GYN and pediatrics providers reported a lack of knowledge about mHealth technologies at rates that exceeded all other specialties (21/27, 77.8% and 37/47, 78.7% vs 45%-71.4% [13/27, 48.1% in Cancer Care and Research; 20/29, 69% in Cardiology; 30/42, 71.4% in Physical and Occupational Therapy; 13/23, 56.5% in Psychiatry and Psychology; 27/60, 45% in Surgery; and 110/214, 51.4% in other specialties], P=.03 and .01 vs family medicine [36/64, 56.3%]). Interestingly, the only self-perceived barrier that was modified by years of experience was the lack of reliable internet access (P=.02 for the overall effect). With

XSL•FO
RenderX

the exception of relatively new providers (0-5 years of experience; 22/95, 23.2% of these providers reported this concern), providers in age groups with ≤20 years of experience (31/73, 42.5% with 6-10 years of experience; 24/64, 37.5% with 11-15 years; and 33/81, 40.7% with 16-20 years) reported higher rates of this concern than those in age groups with >20 years of experience (15/72, 20.8% with 21-25 years of experience; 18/64, 28.1% with 26-30 years; and 23/84, 27.4% with >30 years).

Clinicians' Perceived Barriers Category 2—Patient

Survey respondents reported substantial perceived patient concerns relating to mHealth technologies being too complicated (371/556, 66.7%), lack of access to mHealth technologies (327/556, 58.8%), poor delivery mechanisms (eg, cell service or internet coverage, 252/556, 45.3%), and privacy concerns (207/556, 37.2%). These perceptions did not differ meaningfully by provider type, specialty, or years of experience, with the exception that privacy concerns were more prevalent in doctoral-level providers (105/237, 44.3% vs 93/300, 31%; *P*=.002).

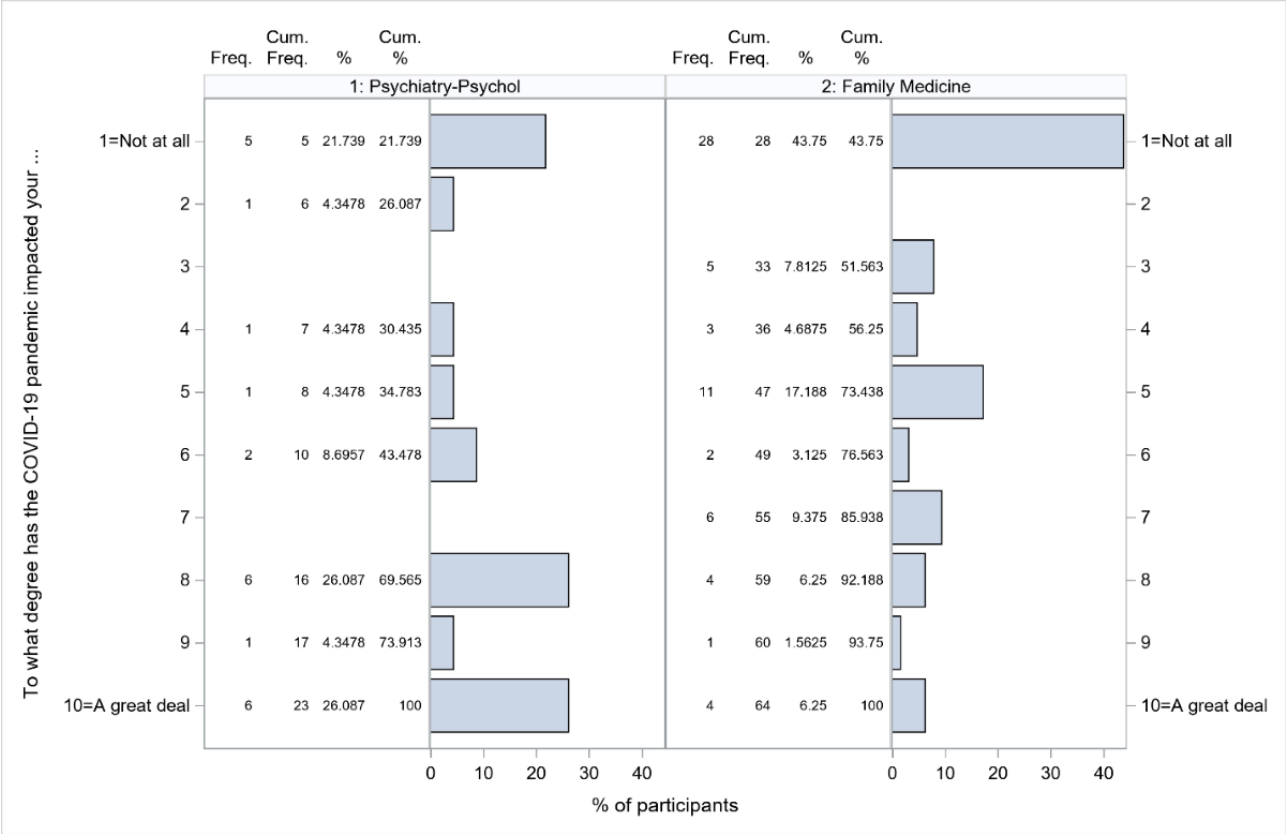
Clinicians' Perceived Barriers Category 3—Organizational

The most prevalent organizational barriers perceived by providers were concerns related to confidentiality (313/556, 56.3%) and that mHealth technologies were too complicated for patients (288/556, 51.8%). Confidentiality concerns differed meaningfully by provider type (149/237, 62.9% and 154/300, 51.3% for doctoral-level vs other providers, *P*=.007), specialty (*P*<.001 for the overall specialty effect; 37/47, 78.7% vs 36/64, 56.3% for pediatrics vs family medicine, *P*=.01), and years of experience (*P*=.03 for the overall effect; no specific trend across age groups). Privacy concerns (168/556, 30.2% prevalence) varied only by years of experience (*P*=.006 for the overall effect; no specific trend across age groups).

COVID-19 and Anticipated mHealth Adoption

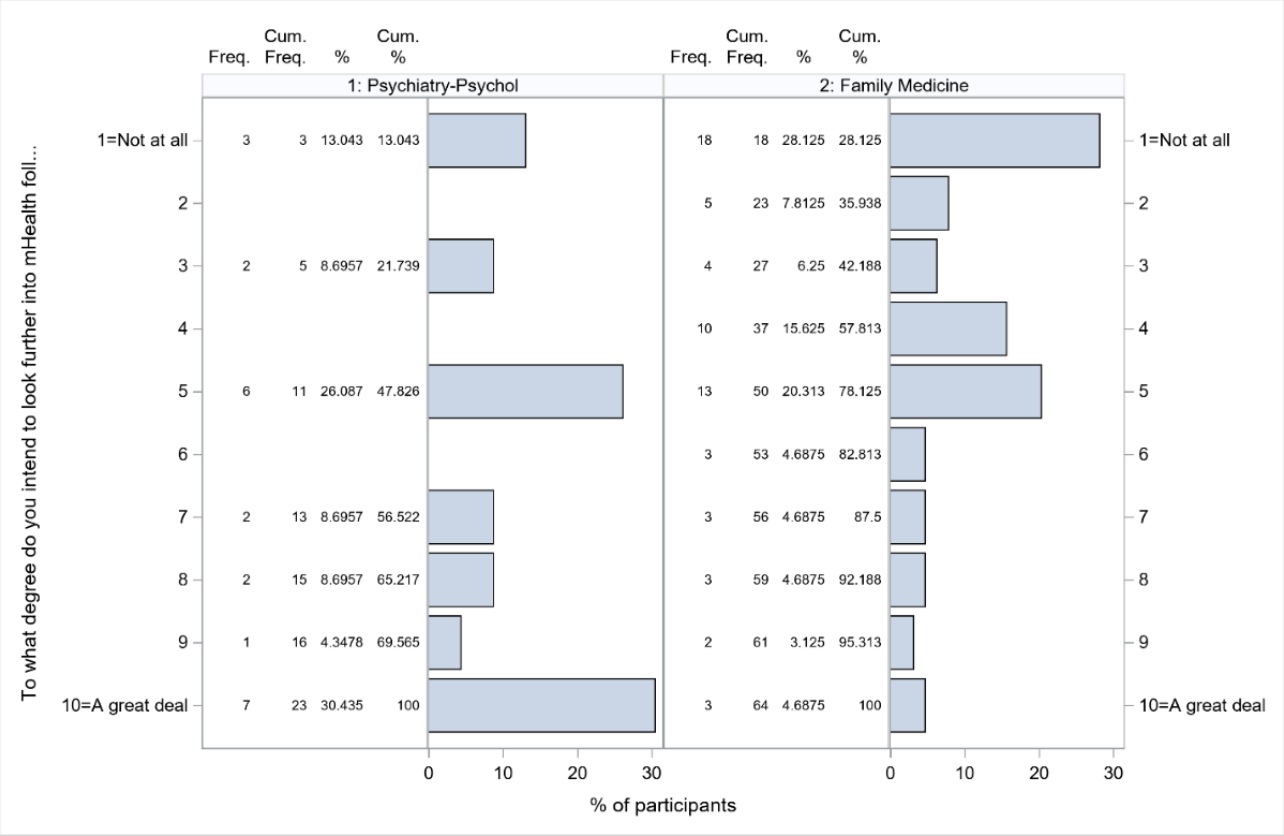
When providers were asked to what degree (1) the COVID-19 pandemic impacted their mHealth adoption and (2) they intend to look further into mHealth following the resumption of normal MCHS activities, meaningful differences were detected only between provider specialties (*P*=.02 and .001 for the overall specialty effects, respectively). These differences were driven by psychiatry and psychology providers, who reported higher scores (10-point Likert scale, where 1=not at all and 10=a great deal) on both survey questions (*P*=.002 and .002 vs family medicine; Figures 2 and 3).

Figure 2. Response distributions for psychiatry and psychology and family medicine for the question "To what degree has the COVID-19 pandemic impacted your mHealth adoption?". mHealth: mobile health.



XSL•FO
RenderX

Figure 3. Response distributions for psychiatry and psychology and family medicine for the question "To what degree do you intend to look further into mHealth following the resumption of normal MCHS activities?" MCHS: Marshfield Clinic Health System; mHealth: mobile health.



Discussion

Principal Findings

The varied responses and rate of mHealth use across provider roles and specialties emphasize the variety and task-specific role mHealth can have in a health system. Some specialties, such as psychology and psychiatry, showed high rates of adoption for specific tasks such as mood and sleep tracking; however, no other specialties reported substantially greater mHealth adoption compared with the reference to family medicine. If looking to increase mHealth use across a health system, leadership should consider identifying specific tasks or poorly performing metrics that mHealth could potentially improve upon.

Our survey grouped potential barriers into 3 levels (provider, patient, and organizational) in line with past qualitative findings [13]. Providers' self-barriers encompassed themes of lack of knowledge about mHealth technologies, lack of time, and lack of patients' access to the internet. Commonly reported barriers relating to both patients and the organization were mHealth technologies being too complicated and concerns related to privacy. The predominant organizational barrier was confidentiality concerns, whereas lack of access (cell phone coverage, internet, and mHealth technology in general) was a frequently perceived barrier for patients.

It is understandable that health care providers feel overwhelmed, with the top barrier to mHealth use being a lack of time and information. These technologies evolve at incredible speeds. How might one stay abreast of the scientific and technological advances of mHealth technologies? Even in the peer-reviewed literature, which can take months or years to publish, we witness an overwhelming ocean of information. At the time of this writing, a Google Scholar search of "mHealth" papers since 2020 (January 1, 2020, to April 5, 2024) yielded nearly 25,000 results and nearly 8000 results since January 1, 2024.

The pace of emerging and simultaneously retiring technologies remains a substantial barrier across many mHealth studies. A typical full-scale trial to evaluate a mHealth initiative lasts more than 5 years from recruitment, during which time many changes within the pertinent technologies will occur or be superseded entirely [20]. Consequentially, many trials are reduced in scope, hindering true evaluation and understanding of the prospect's long-term value.

Years of technological ambition surrounding deep machine learning and voluminous data sets reached fruition in the 2020s with the advent of widely accessible artificial intelligence (AI) apps. These AI-powered breakthroughs have impacted nearly every industry, including medicine, in ways that are still in the infancy of exploration. By digitally processing millions of training samples, including imagery, transcripts, audio recordings, and academic papers, sophisticated computer algorithms have reached new potential in data analysis and user reactivity [21]. What once required thousands of hours and access to prohibitively expensive data centers to compute is now within a finger's reach from any consumer phone or computer.

Leveraging computer-assisted workflows to automate tasks is not a new concept in the medical world. Health care

XSL•FO
RenderX

organizations have spent decades exploring increasingly advanced forms of speech recognition software to facilitate medical transcriptions, among other areas of automation [22]. The latest groundbreaking strides in these efforts come in the form of OpenAI's ChatGPT and associated tool sets [23]. A recent study hypothesized more than 130 different ways ChatGPT could positively benefit both patients and doctors in the foreseeable future, including education, prediction support, prevention of medical errors, record-keeping, and continual clinical assistance [24].

However, many analysts warn that such tool sets—when used in isolation without a human consultant—can yield bad data or other repercussions not yet realized. The most prevalent example of these dangers is how AI modeling is prone to hallucinations, in which the chatbot may return seemingly factual and confident responses but uses nonexistent citations or made-up passages due to anomalies in its training data and other limitations [25,26].

Limitations

The response rate to our mHealth survey was 38% (n/N=916/2410), with a further completion rate of 60.7% (n/N=556/916). It is possible that response bias was present, potentially skewing toward clinicians who have an interest in mHealth technologies. While this response rate is moderately high compared with other surveys of providers, the topic of mHealth being mentioned foremost in the survey invitation may have resulted in an overestimation of mHealth use and intentions.

Notably, MCHS, along with many other health care organizations at the time, was struggling due to the COVID-19 pandemic during our survey timeframe, with rolling temporary furloughs throughout the health system. This limited our possible response rate and created uncertainty in the accuracy and complete capture of our sample. However, our survey was open for 6 weeks with multiple reminder emails sent out, theoretically limiting this effect. Nevertheless, biases in responses may remain due to the work environment and shifting priorities surrounding the COVID-19 pandemic. A future resurveying of providers would help characterize these possible impacts.

Conclusions

Health systems should continue to evaluate mHealth adoption, and more formally and proactively investigate innovative solutions. Consulting with patient safety and legal departments regarding the use of mHealth apps is crucial, as quality clinical outcomes are not often in correlation with popularity ratings on app stores [27]. If a mHealth tool is deemed to be a valuable tool for a hospital or health system, leadership should work toward identifying specific options and methods to address health outcomes and work toward simple and concise implementations to improve adoption and patient outcomes. The American Medical Association provides occasional reports and guidelines surrounding mHealth best practices, but does not have an official lobbying body, with more focus on telehealth [28-30]. The US Department of Health and Human Services provides resources for mHealth developers; however, these are primarily focused on privacy and confidentiality and are of little relevance to providers [31].

This study is arguably a foundational and necessary step in assessing a health system's status and potential for mHealth adoption. Further research and continued partnership with advisors and stakeholders will be needed if the health system hopes to more formally integrate mHealth technologies into rural health care.

Acknowledgments

The authors thank Marie Fleisner for her assistance with editing and formatting the manuscript. The authors also thank the clinicians who participated in this study and the stakeholders who helped and supported them throughout the early stages of this project, stating that their insight and guidance have been a tremendous help. The authors especially thank the project's advisory panel: Kori Krueger, MD; Kelsie-Marie Offenwanger, PsyD; Chris Meyer; John Welch; Peter Shave, MBA; Joe Vechinski, MBA; Ryan Cournoyer, RN, MBA; Jeremy Pomeroy, PhD; Casper Bendixsen, PhD; and Michelle Hamland, DNP, RN, RN-BC.

Conflicts of Interest

None declared.

References

1. Park YT. Emerging new era of mobile health technologies. Healthc Inform Res 2016;22(4):253-254 [FREE Full text] [doi: 10.4258/hir.2016.22.4.253] [Medline: 27895955]
2. Alam MZ, Proteek SM, Hoque I. A systematic literature review on mHealth related research during the COVID-19 outbreak. Health Educ 2023;123(1):19-40 [FREE Full text] [doi: 10.1108/he-08-2022-0067]
3. Alam MMD, Alam MZ, Rahman SA, Taghizadeh SK. Factors influencing mHealth adoption and its impact on mental well-being during COVID-19 pandemic: a SEM-ANN approach. J Biomed Inform 2021;116:103722 [FREE Full text] [doi: 10.1016/j.jbi.2021.103722] [Medline: 33705856]
4. Gagnon MP, Ngangue P, Payne-Gagnon J, Desmartis M. m-Health adoption by healthcare professionals: a systematic review. J Am Med Inform Assoc 2016;23(1):212-220 [FREE Full text] [doi: 10.1093/jamia/ocv052] [Medline: 26078410]
5. Market share of mobile operating systems worldwide from 2009 to 2023, by quarter. Statista. 2023. URL: https://www.statista.com/statistics/272698/global-market-share-held-by-mobile-operating-systems-since-2009 [accessed 2024-04-05]

XSL•FO
RenderX

6. Network coverage outlook and mobility report. Ericsson. 2023. URL: https://www.ericsson.com/en/reports-and-papers/mobility-report/dataforecasts/network-coverage [accessed 2024-04-05]
7. Govoruhina A, Nikiforova A. Digital health shopping assistant with react native: a simple technological solution to a complex health problem. 2022 Presented at: 2022 International Conference on Intelligent Data Science Technologies and Applications (IDSTA); September 05-07, 2022; San Antonio, TX, USA. [doi: 10.1109/idsta55301.2022.9923047]
8. Fentaw AE. Cross platform mobile application development: a comparison study of react native vs flutter. JYX Digital Repository. 2020. URL: https://jyx.jyu.fi/handle/123456789/70969# [accessed 2024-05-09]
9. Number of connected wearable devices worldwide from 2016 to 2022. Statista. 2023. URL: https://www.statista.com/statistics/487291/global-connected-wearable-devices/ [accessed 2024-04-05]
10. Niwa M. mHealth as a component of next-generation health care. In: Kodama K, Sengoku S, editors. Mobile Health (mHealth): Rethinking Innovation Management to Harmonize AI and Social Design. Singapore: Springer Nature; 2022:189-209.
11. Non-invasive blood glucose monitoring closer as Apple hits major milestone. Diabaetes.co.uk. 2023. URL: https://www.diabetes.co.uk/news/2023/mar/non-invasive-blood-glucose-monitoring-closer-as-apple-hits-major-milestone.html [accessed 2023-03-30]
12. Sheng A, Lin L, Zhu J, Zhuang J, Li J, Chang L, et al. Micro/nanodevices for assessment and treatment in stomatology and ophthalmology. Microsyst Nanoeng 2021;7:11 [FREE Full text] [doi: 10.1038/s41378-021-00238-1] [Medline: 33532080]
13. Khatun F, Heywood AE, Ray PK, Hanifi SM, Bhuiya A, Liaw S. Determinants of readiness to adopt mHealth in a rural community of Bangladesh. Int J Med Inform 2015 Oct;84(10):847-856. [doi: 10.1016/j.ijmedinf.2015.06.008] [Medline: 26194141]
14. Weichelt B, Bendixsen C, Patrick T. A model for assessing necessary conditions for rural health care's mobile health readiness: qualitative assessment of clinician-perceived barriers. JMIR Mhealth Uhealth 2019;7(11):e11915 [FREE Full text] [doi: 10.2196/11915] [Medline: 31702564]
15. Shimpi N, Glurich I, Maybury C, Wang MQ, Hashimoto K, Acharya A, et al. Knowledge, attitudes, behaviors of women related to pregnancy, and early childhood caries prevention: a cross-sectional pilot study. J Prim Care Community Health 2021;12:21501327211013302 [FREE Full text] [doi: 10.1177/21501327211013302] [Medline: 33949227]
16. About marshfield clinic health system. Marshfield Clinic Health System. 2023. URL: https://www.marshfieldclinic.org/about-us [accessed 2023-09-07]
17. McNutt LA, Wu C, Xue X, Hafner JP. Estimating the relative risk in cohort studies and clinical trials of common outcomes. Am J Epidemiol 2003;157(10):940-943 [FREE Full text] [doi: 10.1093/aje/kwg074] [Medline: 12746247]
18. Kruskal WH, Wallis WA. Use of ranks in one-criterion variance analysis. J Am Stat Assoc 1952;47(260):583-621 [FREE Full text] [doi: 10.1080/01621459.1952.10483441]
19. Hollander M, Wolfe DA. Nonparametric Statistical Methods. Hoboken, New Jersey: Wiley; 1999.
20. Kumar S, Nilsen WJ, Abernethy A, Atienza A, Patrick K, Pavel M, et al. Mobile health technology evaluation: the mHealth evidence workshop. Am J Prev Med 2013;45(2):228-236 [FREE Full text] [doi: 10.1016/j.amepre.2013.03.017] [Medline: 23867031]
21. Glock P, von Alemann S. The paradigm shift in AI: from human labor to humane creativity. In: Waltl B, Schindler D, Jacob K, Strathausen R, editors. Liquid Legal – Humanization and the Law. Singapore: Springer International Publishing; 2022:215-230.
22. Borowitz SM. Computer-based speech recognition as a replacement for medical transcription. Pediatr Res 1999;45(7):120 [FREE Full text] [doi: 10.1203/00006450-199904020-00715]
23. Hughes A. ChatGPT: everything you need to know about OpenAI's GPT-4 tool. BBC Science Focus Magazine. 2023. URL: https://www.sciencefocus.com/future-technology/gpt-3 [accessed 2024-05-09]
24. Javaid M, Haleem A, Singh RP. ChatGPT for healthcare services: an emerging stage for an innovative perspective. BenchCouncil Trans Benchmarks Stand Eval 2023;3(1):100105 [FREE Full text] [doi: 10.1016/j.tbench.2023.100105]
25. Athaluri SA, Manthena SV, Kesapragada VSRKM, Yarlagadda V, Dave T, Duddumpudi RTS. Exploring the boundaries of reality: investigating the phenomenon of artificial intelligence hallucination in scientific writing through ChatGPT references. Cureus 2023;15(4):e37432 [FREE Full text] [doi: 10.7759/cureus.37432] [Medline: 37182055]
26. Brereton D. Bing AI can't be trusted. DKB Blog. 2023. URL: https://dkb.blog/p/bing-ai-cant-be-trusted [accessed 2023-05-25]
27. Wyatt JC. How can clinicians, specialty societies and others evaluate and improve the quality of apps for patient use? BMC Med 2018;16(1):225 [FREE Full text] [doi: 10.1186/s12916-018-1211-7] [Medline: 30501638]
28. AMA adopts principles to promote safe, effective mHealth applications. American Medical Association. 2016. URL: https://www.ama-assn.org/press-center/press-releases/ama-adopts-principles-promote-safe-effective-mhealth-applications [accessed 2024-04-05]
29. AMA's advocacy efforts push for permanent telehealth advancements. American Medical Association. 2022. URL: https://www.ama-assn.org/practice-management/digital/amas-advocacy-efforts-push-permanent-telehealth-advancements [accessed 2024-05-09]
30. Comprehensive telehealth reform is critical to the future of health care. American Medical Association. 2023. URL: https://www.ama-assn.org/amaone/ama-recovery-plan-america-s-physicians-supporting-telehealth [accessed 2024-04-05]

XSL•FO
RenderX

31. Resources for mobile health apps developers. US Department of Health and Human Services. 2022. URL: https://www.hhs.gov/hipaa/for-professionals/special-topics/health-apps/index.html [accessed 2024-04-04]

Abbreviations

AI: artificial intelligence
MCHS: Marshfield Clinic Health System
mHealth: mobile health

Edited by C Jacob; submitted 13.12.23; peer-reviewed by M Niwa, D Singh; comments to author 22.01.24; revised version received 05.04.24; accepted 08.04.24; published 24.06.24.

Please cite as:
Weichelt BP, Burke R, Kieke B, Pilz M, Shimpi N
Provider Adoption of mHealth in Rural Patient Care: Web-Based Survey Study
JMIR Hum Factors 2024;11:e55443
URL: https://humanfactors.jmir.org/2024/1/e55443
doi:10.2196/55443
PMID:

©Bryan P Weichelt, Rick Burke, Burney Kieke, Matt Pilz, Neel Shimpi. Originally published in JMIR Human Factors (https://humanfactors.jmir.org), 24.06.2024. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on https://humanfactors.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Leveraging Generative AI Tools to Support the Development of Digital Solutions in Health Care Research: Case Study

Danissa V Rodriguez[1], PhD; Katharine Lawrence[1,2], MPH, MD; Javier Gonzalez[2], BS; Beatrix Brandfield-Harvey[1], BA; Lynn Xu[1], MPH; Sumaiya Tasneem[1], MPH; Defne L Levine[1], MPH; Devin Mann[1,2], MS, MD

[1]Department of Population Health, New York University Grossman School of Medicine, New York, NY, United States

[2]Medical Center Information Technology, Department of Health Informatics, New York University Langone Health, New York, NY, United States

Corresponding Author:
Danissa V Rodriguez, PhD
Department of Population Health
New York University Grossman School of Medicine
227 East 30th Street
6th Floor
New York, NY, 10016
United States
Phone: 1 646 501 2684
Email: danissa.rodriguez@nyulangone.org

Abstract

Background: Generative artificial intelligence has the potential to revolutionize health technology product development by improving coding quality, efficiency, documentation, quality assessment and review, and troubleshooting.

Objective: This paper explores the application of a commercially available generative artificial intelligence tool (ChatGPT) to the development of a digital health behavior change intervention designed to support patient engagement in a commercial digital diabetes prevention program.

Methods: We examined the capacity, advantages, and limitations of ChatGPT to support digital product idea conceptualization, intervention content development, and the software engineering process, including software requirement generation, software design, and code production. In total, 11 evaluators, each with at least 10 years of experience in fields of study ranging from medicine and implementation science to computer science, participated in the output review process (ChatGPT vs human-generated output). All had familiarity or prior exposure to the original personalized automatic messaging system intervention. The evaluators rated the ChatGPT-produced outputs in terms of understandability, usability, novelty, relevance, completeness, and efficiency.

Results: Most metrics received positive scores. We identified that ChatGPT can (1) support developers to achieve high-quality products faster and (2) facilitate nontechnical communication and system understanding between technical and nontechnical team members around the development goal of rapid and easy-to-build computational solutions for medical technologies.

Conclusions: ChatGPT can serve as a usable facilitator for researchers engaging in the software development life cycle, from product conceptualization to feature identification and user story development to code generation.

Trial Registration: ClinicalTrials.gov NCT04049500; https://clinicaltrials.gov/ct2/show/NCT04049500

(JMIR Hum Factors 2024;11:e52885) doi:10.2196/52885

KEYWORDS

digital health; GenAI; generative; artificial intelligence; ChatGPT; software engineering; mHealth; mobile health; app; apps; application; applications; diabetes; diabetic; diabetes prevention; digital prescription; software; engagement; behaviour change; behavior change; developer; developers; LLM; LLMs; language model; language models; NLP; natural language processing

Introduction

Health care has undergone a digital transformation, resulting in a growing reliance on software engineering for medical use cases, including health care research. However, little guidance exists for health researchers on how to effectively develop digital health interventions [1]; in particular, software development challenges that include expertise gaps in coding, custom development needs, high costs, and time constraints result in

XSL•FO
RenderX

multilevel barriers to designing and deploying a usable, scalable, and sustainable digital health product [1].

Generative artificial intelligence (GenAI) technologies such as ChatGPT can potentially support researchers in health technology endeavors by providing foundational frameworks and processes for the software development life cycle [2]. These systems can help reduce time and enhance precision for technology-based research projects by supporting both nonprogrammers and experienced programmers in code development, troubleshooting, and cleaning [2]. Moreover, the ability to use GenAI to generate content from different perspectives (expert or nonexpert) can facilitate and improve communication between technical and nontechnical team members of multidisciplinary teams. For example, a nontechnical team member can write their ideas in natural text and then use GenAI to request assistance in creating discussion points to communicate to a technical team audience. GenAI tools may also help health technology researchers refine research questions, identify appropriate theoretical frameworks and models, and leverage popular implementation strategies such as design thinking to build effective, theory-grounded, and evidence-based digital health interventions. ChatGPT (OpenAI, Microsoft Corporation) has already demonstrated feasibility as a support tool for clinical decision support development in health care [3], and more broadly as a coding copilot in programming and engineering [4,5].

This study explores the use of ChatGPT to recreate a personalized automatic messaging system (PAMS), which was developed as part of a digital health research initiative to support patient engagement with a commercial digital diabetes prevention program (dDPP). We examine the capacity, advantages, and limitations of ChatGPT to support product ideation and conceptualization, intervention content development, and the software engineering process including software requirement generation, software design, and code production. This paper provides insights to support the GenAI-assisted development of computational tools that are usable, reliable, extensible, and in line with the standards of modern coding practices. The framework includes prompts for both the intervention conceptualization as well as the main phases of the software development process.

Methods

Settings and Intervention Development Context

In previous work [6], we described the development of PAMS, a novel integrated multicomponent communications platform, to promote patient-provider communication and patient engagement in a commercial dDPP (Noom; Noom, Inc). The PAMS intervention included early prototyping and user testing, a technical development phase, and a randomized controlled trial. The core content and user experience features of PAMS were identified, prototyped, and evaluated using the well-established design thinking "discover, define, design, and test" approach to iteratively gather information, define, design, and refine the engagement intervention [7]. Stakeholders included: patients with prediabetes and their support network (eg, caregivers and partners), primary care providers, health technologists, programmers and computer scientists, behavioral change theorists and subject matter experts, the research administrative team, and dDPP product developers and coaches. The main components of this PAMS intervention include (1) a theory-driven behavior change messaging library, (2) a personalized automated message system delivery platform (SMS text messaging–based), and (3) EHR-integrated data visualizations. The PAMS messaging library uses an integrated framework that combines established theoretical models for behavior change with human-centered design strategies to maximize the evidence-based conditions for behavior change and the user acceptance and use of a digital health product. The technical development of PAMS followed an agile software development approach based on incremental 2-week sprint cycles consisting of requirement planning, design, development, and testing of a specific set of functional features. In this paper, we will recreate this development process using GenAI (ChatGPT).

ChatGPT-PAMS Experiment Design

To evaluate the effectiveness of using GenAI to support the development of digital tools in medical settings, our experiment is based on recreating PAMS using GenAI (ChatGPT) and evaluating human-generated vs ChatGPT-generated documentation. To accurately capture the ideation and development process, our multidisciplinary team reviewed all documentation and processes used in the early stages of PAMS conceptualization, including supporting theoretical models, content and features, and technical development. We then recreated these processes via a series of prompts for ChatGPT-4 to assist with the generation of theory, content, user stories, requirement documents, design diagrams, and the code for a subset of the requirements. Outputs from ChatGPT were reviewed and compared to human-generated documentation by 11 evaluating team members. Evaluators consisted of clinicians, behavioral scientists, programmers, and research staff working in digital health and technology for behavior change research. Collectively, they represent more than 50 years of clinical, research, design, and computer science experience. The evaluators independently rated the quality of various aspects of information provided by ChatGPT on a Likert scale, where higher ratings indicated greater quality of information (*1: very poor; 2: poor; 3: acceptable; 4: good; 5: very good; N/A: not applicable*). Aspects of evaluation included: understandability (Does this output make sense given the context of the study and prompts?), novelty (Were new ideas generated?) [3], usability (Does this create a usable output?), relevance (Does this create a useful output?), efficiency (Would having these outputs have saved time?), and potential for bias (What unintended consequences might arise from these outputs?) [6]. Evaluators were also asked to give an overall score on the quality of the ChatGPT output (Overall, how good would you say this output is?). Post review, a group debrief was conducted, using a semistructured interview guide to facilitate discussion regarding perceptions of outputs and rationale for ratings.

Ethical Considerations

Ethical considerations helped guide the initial development of research methods and reduce potential risks for participants in

XSL•FO
RenderX

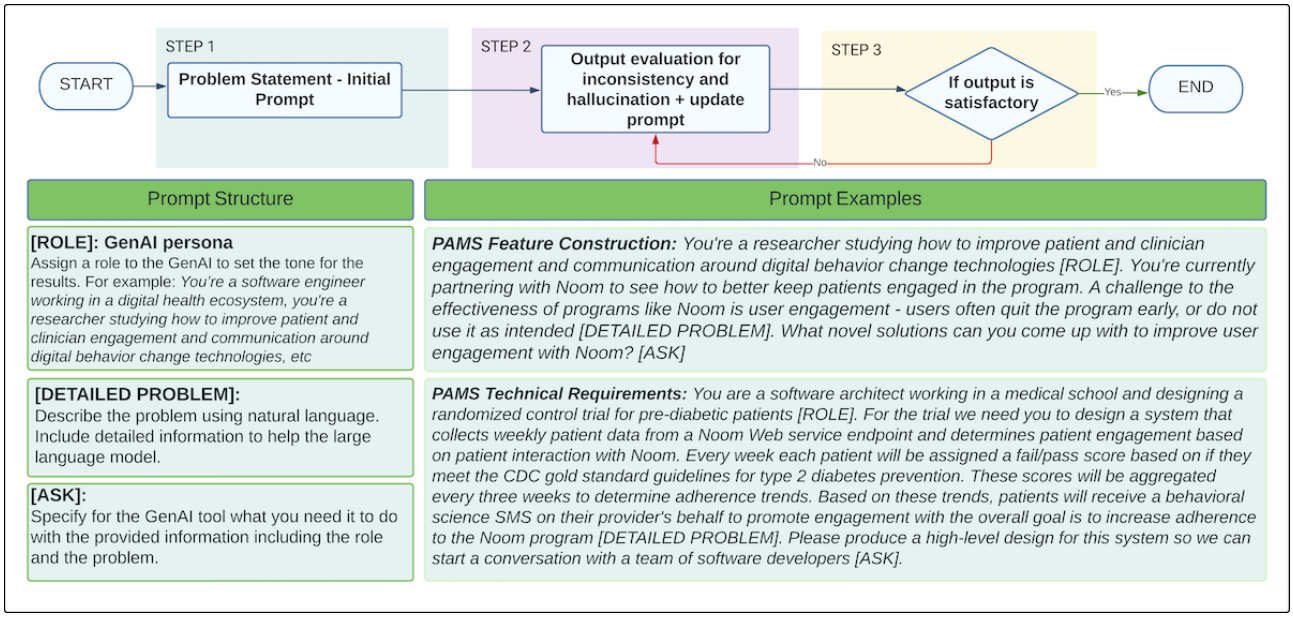
the original study implementation with the PAMS intervention [7]. Recreating the technical development of a system previously built as part of the dDPP randomized controlled trial (NCT04049500) has not introduced any new risks to patients. Patients were not involved in this research examining the use of GenAI in the development of digital health care solutions. No patient data was used in the prompt generation phase.

Regarding ethical considerations for the methods used in this paper, as an attempt to mitigate evaluator biases, we worked with a diverse team of evaluators who were aware of the initial study but were not necessarily involved in the technical development. Additionally, we understand the limitations and concerns of the use of ChatGPT including possible hallucinations and incorrect answers. Thus, we emphasize the need for human expertise to identify correct and incorrect outputs and have flagged this as a study consideration. When developing the GenAI-based solution, we used the same considerations for data security, patient usability, accessibility, and data privacy used in the original human-developed solution.

Prompt Generation Framework

Prompt engineering focuses on the skill of designing and creating effective prompts that guide ChatGPT to produce the best possible output for your task. We followed existing literature [8-11] combined with our expertise and experimentation to provide a framework that yields the best result when developing a digital solution like PAMS (Figure 1).

Figure 1. ChatGPT prompt structure and prompt examples. CDC: Centers for Disease Control and Prevention; GenAI: generative artificial intelligence; PAMS: personalized automatic messaging system.



Results

PAMS Concept and User Experience Generation

Overview

Core components of the PAMS intervention were conceptualized and designed via an underlying behavior change theory, design principles and personas, and a message content library.

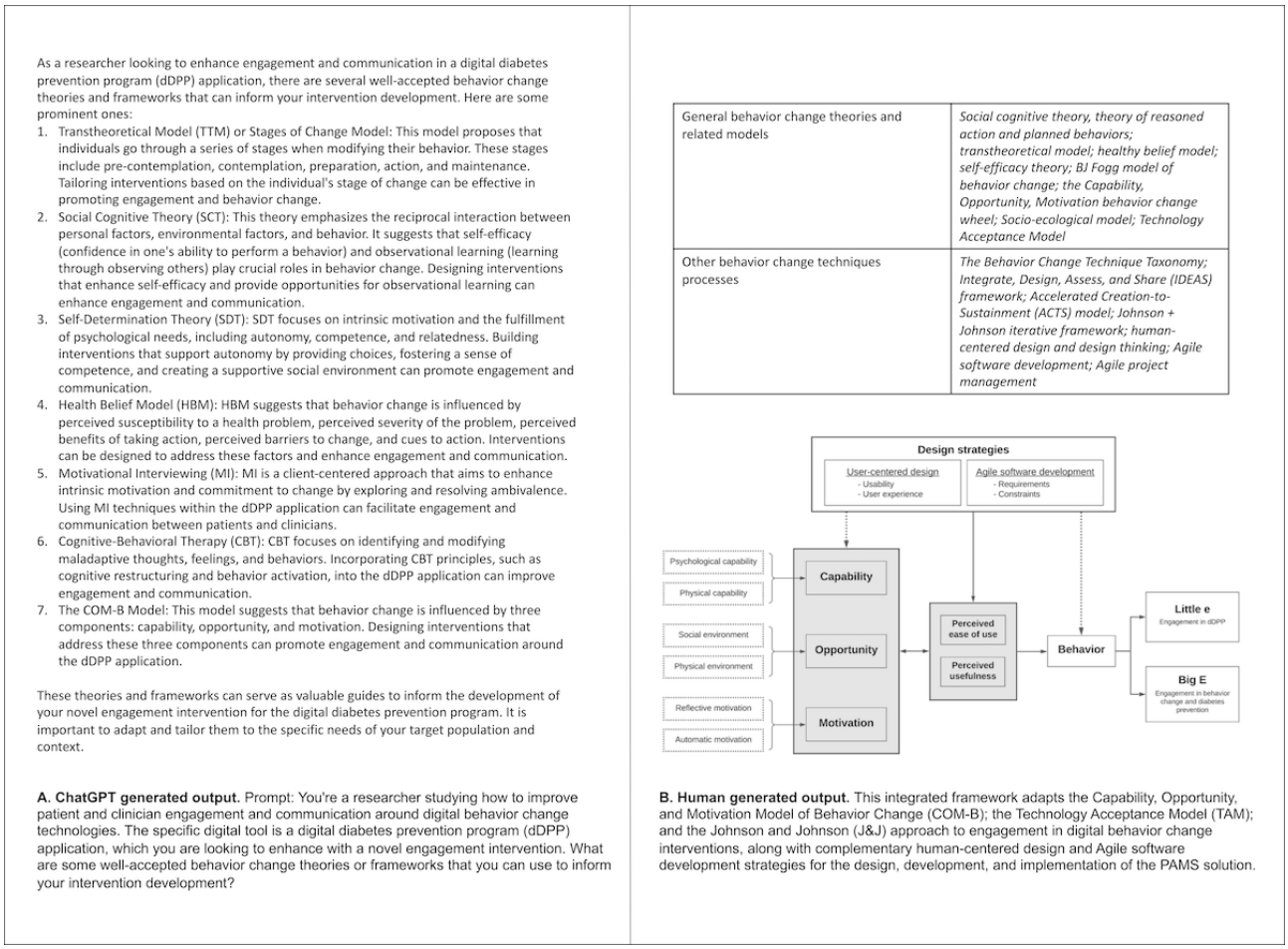
Underpinning Behavior Change Theory and Approach

Human-Generated Solution

Leveraging behavior change literature review and interviews with behavior change theory content experts (n=4), the research team initially identified ten unique behavior change theories and six process models that were considered to be an appropriate fit for the aims of the overall intervention. A unique model was developed that captured (1) the relevant underlying behavior change theory, (2) implementation strategies, and (3) unique contexts of the technology environment (Figure 2A).

XSL•FO
RenderX

Figure 2. Underpinning behavior change theory and approach outcome of ChatGPT vs human-generated output. ACTS: Accelerated Creation to Sustainment; BJ: Brian Jeffrey; CBT: cognitive behavioral therapy; COM-B: capability, opportunity, and motivation model of behavior change; dDPP: digital diabetes prevention program; HBM: health belief model; IDEAS: Integrate, Design, Assess, and Share; J&J: Johnson and Johnson; MI: motivational interviewing; SCT: social cognitive theory; SDT: self-determination theory; TTM: transtheoretical model.



GenAI Solution

When prompted, ChatGPT identified seven relevant well-accepted behavior change theories and frameworks to inform a "dDPP support intervention" (Figure 2B). It did not provide information on the transtheoretical domains framework, or the taxonomy of behavior change techniques, but when prompted on these, identified both as acceptable strategies for use.

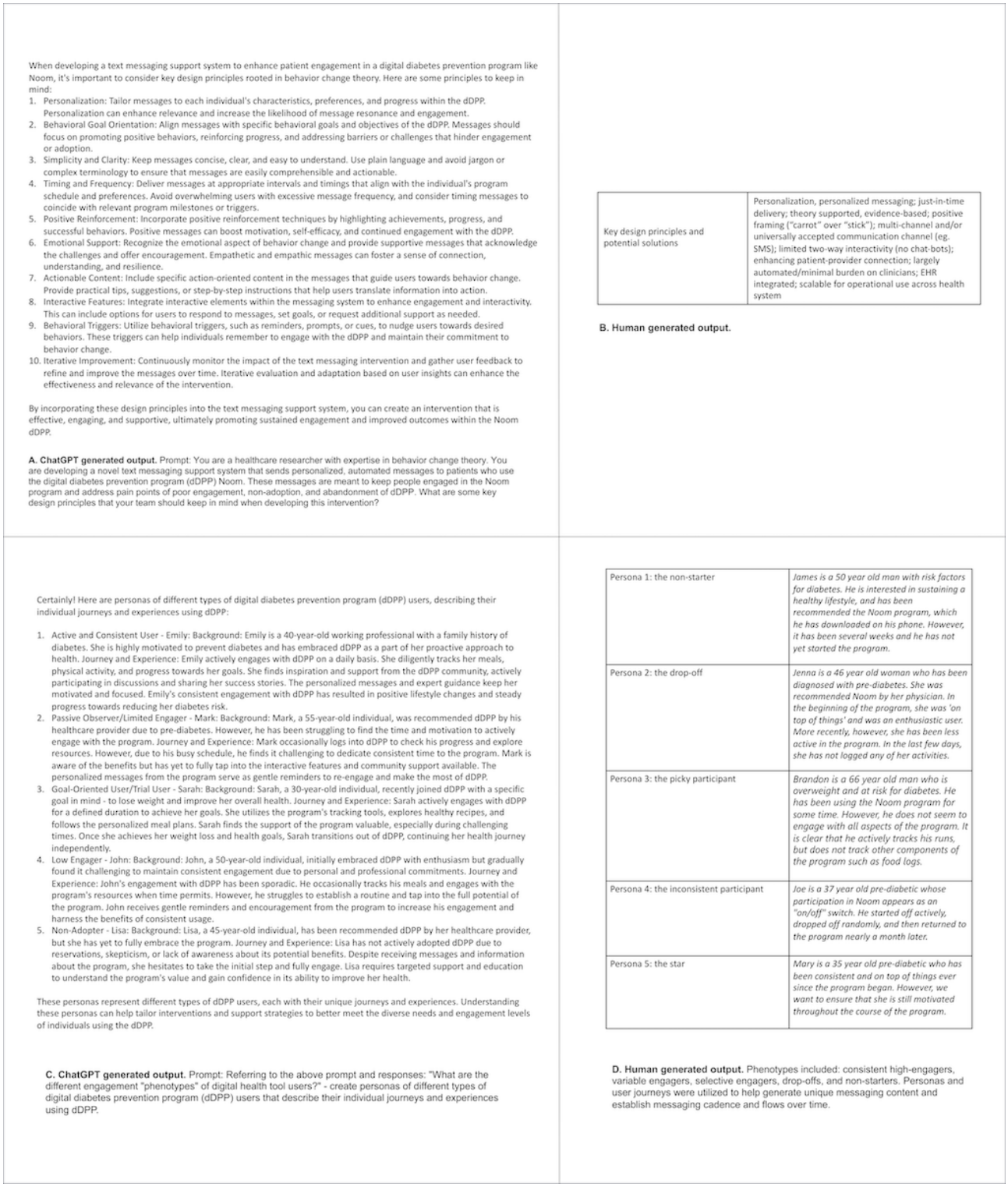
User Experience: Design Principles, Personas, and Messaging Content

Human-Generated Solution

The research team used a human-centered design approach to identify key design principles, defined as the set of considerations that form the basis of the PAMS product (Figure 3B). These were developed from insights gathered via a review of relevant digital behavior change research, consultation with content and theoretical experts in digital health and implementation science (n=3), and two group interviews (n=9). From these insights, five relevant fictional personas were designed to capture the various phenotypes of user engagement with the commercial dDPP, along with unique user journeys developed to describe their projected engagement with the program over time (Figure 3D). Overall, over 193 unique messages were developed, each grounded by a relevant behavior change technique and tailored to an individual phenotype's user journey. These elements were continuously revisited and refined during the testing phases of the dDPP research. This included a 6-month near-live user testing phase consisting of nine patients engaging with various iterations of the PAMS prototype, and a 12-month live single-arm pilot phase consisting of 25 patients using PAMS-beta with the commercial dDPP platform.

XSL•FO
RenderX

Figure 3. User experience: design principles, personas, and messaging content outcome of ChatGPT vs human-generated output. dDPP: digital diabetes prevention program; EHR: electronic health record.



When developing a text messaging support system to enhance patient engagement in a digital diabetes prevention program like Noom, it's important to consider key design principles rooted in behavior change theory. Here are some principles to keep in mind:

1. Personalization: Tailor messages to each individual's characteristics, preferences, and progress within the dDPP. Personalization can enhance relevance and increase the likelihood of message resonance and engagement.
2. Behavioral Goal Orientation: Align messages with specific behavioral goals and objectives of the dDPP. Messages should focus on promoting positive behaviors, reinforcing progress, and addressing barriers or challenges that hinder engagement or adoption.
3. Simplicity and Clarity: Keep messages concise, clear, and easy to understand. Use plain language and avoid jargon or complex terminology to ensure that messages are easily comprehensible and actionable.
4. Timing and Frequency: Deliver messages at appropriate intervals and timings that align with the individual's program schedule and preferences. Avoid overwhelming users with excessive message frequency, and consider timing messages to coincide with relevant program milestones or triggers.
5. Positive Reinforcement: Incorporate positive reinforcement techniques by highlighting achievements, progress, and successful behaviors. Positive messages can boost motivation, self-efficacy, and continued engagement with the dDPP.
6. Emotional Support: Recognize the emotional aspect of behavior change and provide supportive messages that acknowledge the challenges and offer encouragement. Empathetic and empathic messages can foster a sense of connection, understanding, and resilience.
7. Actionable Content: Include specific action-oriented content in the messages that guide users towards behavior change. Provide practical tips, suggestions, or step-by-step instructions that help users translate information into action.
8. Interactive Features: Integrate interactive elements within the messaging system to enhance engagement and interactivity. This can include options for users to respond to messages, set goals, or request additional support as needed.
9. Behavioral Triggers: Utilize behavioral triggers, such as reminders, prompts, or cues, to nudge users towards desired behaviors. These triggers can help individuals remember to engage with the dDPP and maintain their commitment to behavior change.
10. Iterative Improvement: Continuously monitor the impact of the text messaging intervention and gather user feedback to refine and improve the messages over time. Iterative evaluation and adaptation based on user insights can enhance the effectiveness and relevance of the intervention.

By incorporating these design principles into the text messaging support system, you can create an intervention that is effective, engaging, and supportive, ultimately promoting sustained engagement and improved outcomes within the Noom dDPP.

A. ChatGPT generated output. Prompt: You are a healthcare researcher with expertise in behavior change theory. You are developing a novel text messaging support system that sends personalized, automated messages to patients who use the digital diabetes prevention program (dDPP) Noom. These messages are meant to keep people engaged in the Noom program and address pain points of poor engagement, non-adoption, and abandonment of dDPP. What are some key design principles that your team should keep in mind when developing this intervention?

Key design principles and potential solutions	Personalization, personalized messaging; just-in-time delivery; theory supported, evidence-based; positive framing ("carrot" over "stick"); multi-channel and/or universally accepted communication channel (eg. SMS); limited two-way interactivity (no chat-bots); enhancing patient-provider connection; largely automated/minimal burden on clinicians; EHR integrated; scalable for operational use across health system

B. Human generated output.

Certainly! Here are personas of different types of digital diabetes prevention program (dDPP) users, describing their individual journeys and experiences using dDPP:

1. Active and Consistent User - Emily: Background: Emily is a 40-year-old working professional with a family history of diabetes. She is highly motivated to prevent diabetes and has embraced dDPP as a part of her proactive approach to health. Journey and Experience: Emily actively engages with dDPP on a daily basis. She diligently tracks her meals, physical activity, and progress towards her goals. She finds inspiration and support from the dDPP community, actively participating in discussions and sharing her success stories. The personalized messages and expert guidance keep her motivated and focused. Emily's consistent engagement with dDPP has resulted in positive lifestyle changes and steady progress towards reducing her diabetes risk.
2. Passive Observer/Limited Engager - Mark: Background: Mark, a 55-year-old individual, was recommended dDPP by his healthcare provider due to pre-diabetes. However, he has been struggling to find the time and motivation to actively engage with the program. Journey and Experience: Mark occasionally logs into dDPP to check his progress and explore resources. However, due to his busy schedule, he finds it challenging to dedicate consistent time to the program. Mark is aware of the benefits but has yet to fully tap into the interactive features and community support available. The personalized messages from the program serve as gentle reminders to re-engage and make the most of dDPP.
3. Goal-Oriented User/Trial User - Sarah: Background: Sarah, a 30-year-old individual, recently joined dDPP with a specific goal in mind - to lose weight and improve her overall health. Journey and Experience: Sarah actively engages with dDPP for a defined duration to achieve her goals. She utilizes the program's tracking tools, explores healthy recipes, and follows the personalized meal plans. Sarah finds the support of the program valuable, especially during challenging times. Once she achieves her weight loss and health goals, Sarah transitions out of dDPP, continuing her health journey independently.
4. Low Engager - John: Background: John, a 50-year-old individual, initially embraced dDPP with enthusiasm but gradually found it challenging to maintain consistent engagement due to personal and professional commitments. Journey and Experience: John's engagement with dDPP has been sporadic. He occasionally tracks his meals and engages with the program's resources when time permits. However, he struggles to establish a routine and tap into the full potential of the program. John receives gentle reminders and encouragement from the program to increase his engagement and harness the benefits of consistent usage.
5. Non-Adopter - Lisa: Background: Lisa, a 45-year-old individual, has been recommended dDPP by her healthcare provider, but she has yet to fully embrace the program. Journey and Experience: Lisa has not actively adopted dDPP due to reservations, skepticism, or lack of awareness about its potential benefits. Despite receiving messages and information about the program, she hesitates to take the initial step and fully engage. Lisa requires targeted support and education to understand the program's value and gain confidence in its ability to improve her health.

These personas represent different types of dDPP users, each with their unique journeys and experiences. Understanding these personas can help tailor interventions and support strategies to better meet the diverse needs and engagement levels of individuals using the dDPP.

C. ChatGPT generated output. Prompt: Referring to the above prompt and responses: "What are the different engagement "phenotypes" of digital health tool users?" - create personas of different types of digital diabetes prevention program (dDPP) users that describe their individual journeys and experiences using dDPP.

Persona 1: the non-starter	*James is a 50 year old man with risk factors for diabetes. He is interested in sustaining a healthy lifestyle, and has been recommended the Noom program, which he has downloaded on his phone. However, it has been several weeks and he has not yet started the program.*
Persona 2: the drop-off	*Jenna is a 46 year old woman who has been diagnosed with pre-diabetes. She was recommended Noom by her physician. In the beginning of the program, she was 'on top of things' and was an enthusiastic user. More recently, however, she has been less active in the program. In the last few days, she has not logged any of her activities.*
Persona 3: the picky participant	*Brandon is a 66 year old man who is overweight and at risk for diabetes. He has been using the Noom program for some time. However, he does not seem to engage with all aspects of the program. It is clear that he actively tracks his runs, but does not track other components of the program such as food logs.*
Persona 4: the inconsistent participant	*Joe is a 37 year old pre-diabetic whose participation in Noom appears as an "on/off" switch. He started off actively, dropped off randomly, and then returned to the program nearly a month later.*
Persona 5: the star	*Mary is a 35 year old pre-diabetic who has been consistent and on top of things ever since the program began. However, we want to ensure that she is still motivated throughout the course of the program.*

D. Human generated output. Phenotypes included: consistent high-engagers, variable engagers, selective engagers, drop-offs, and non-starters. Personas and user journeys were utilized to help generate unique messaging content and establish messaging cadence and flows over time.

GenAI Solution

ChatGPT was prompted from multiple perspectives (researcher, clinician, and patient) to identify key design principles (Figure 3A) and sample solutions for the PAMS intervention. It also provided common engagement phenotypes for digital health tool users, based on patterns of use, frequency, duration, and "other elements." Of note, nonadopters were not identified within the initial round of phenotypes. ChatGPT also developed personas for each of the identified engagement phenotypes, including persona names, backgrounds, and individual journeys. ChatGPT was able to produce five to ten unique messages targeted toward each phenotype and to adapt these messages based on various additional prompts. The user types or personas generated by ChatGPT are consistent with the human-generated users and cover all the phenotypes identified in our previous research (eg, mapping to a specific behavior change technique and reflecting a key design principle; Figure 3C).

XSL•FO
RenderX

PAMS Technical Development

Overview

The technical development includes a PAMS requirements document and architectural design and code.

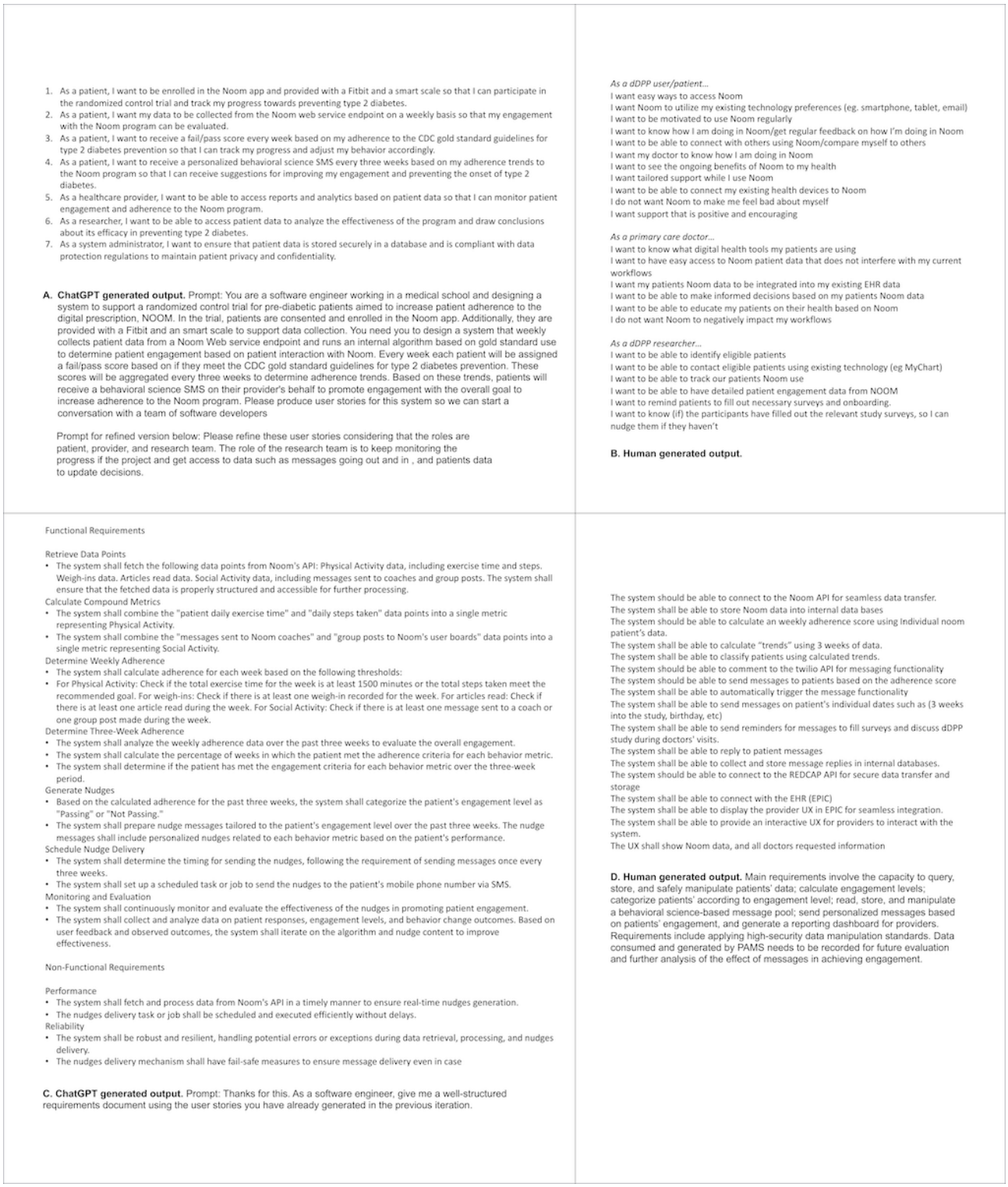
Technical Requirements (User Stories)

Human-Generated Solution

Following the data collection and intervention design period, we created, as a team, a series of user stories (Figure 4B) which were followed by system requirements to describe the intended use cases, features, and challenges of the proposed PAMS software. Initial system requirements represent the "minimum viable product" that was developed, piloted, and further refined (Figure 4D). Our development team followed software engineering principles to generate the requirements document.

XSL•FO
RenderX

Figure 4. User stories and technical requirements outcome of ChatGPT vs human-generated outputs. API: application programming interface; CDC: Centers for Disease Control and Prevention; EHR: electronic health record; PAMS: personalized automatic messaging system; REDCap: Research Electronic Data Capture; UX: user interface.



1. As a patient, I want to be enrolled in the Noom app and provided with a Fitbit and a smart scale so that I can participate in the randomized control trial and track my progress towards preventing type 2 diabetes.
2. As a patient, I want my data to be collected from the Noom web service endpoint on a weekly basis so that my engagement with the Noom program can be evaluated.
3. As a patient, I want to receive a fail/pass score every week based on my adherence to the CDC gold standard guidelines for type 2 diabetes prevention so that I can track my progress and adjust my behavior accordingly.
4. As a patient, I want to receive a personalized behavioral science SMS every three weeks based on my adherence trends to the Noom program so that I can receive suggestions for improving my engagement and preventing the onset of type 2 diabetes.
5. As a healthcare provider, I want to be able to access reports and analytics based on patient data so that I can monitor patient engagement and adherence to the Noom program.
6. As a researcher, I want to be able to access patient data to analyze the effectiveness of the program and draw conclusions about its efficacy in preventing type 2 diabetes.
7. As a system administrator, I want to ensure that patient data is stored securely in a database and is compliant with data protection regulations to maintain patient privacy and confidentiality.

A. ChatGPT generated output. Prompt: You are a software engineer working in a medical school and designing a system to support a randomized control trial for pre-diabetic patients aimed to increase patient adherence to the digital prescription, NOOM. In the trial, patients are consented and enrolled in the Noom app. Additionally, they are provided with a Fitbit and an smart scale to support data collection. You need you to design a system that weekly collects patient data from a Noom Web service endpoint and runs an internal algorithm based on gold standard use to determine patient engagement based on patient interaction with Noom. Every week each patient will be assigned a fail/pass score based on if they meet the CDC gold standard guidelines for type 2 diabetes prevention. These scores will be aggregated every three weeks to determine adherence trends. Based on these trends, patients will receive a behavioral science SMS on their provider's behalf to promote engagement with the overall goal to increase adherence to the Noom program. Please produce user stories for this system so we can start a conversation with a team of software developers

Prompt for refined version below: Please refine these user stories considering that the roles are patient, provider, and research team. The role of the research team is to keep monitoring the progress if the project and get access to data such as messages going out and in , and patients data to update decisions.

As a dDPP user/patient…
I want easy ways to access Noom
I want Noom to utilize my existing technology preferences (eg. smartphone, tablet, email)
I want to be motivated to use Noom regularly
I want to know how I am doing in Noom/get regular feedback on how I'm doing in Noom
I want to be able to connect with others using Noom/compare myself to others
I want my doctor to know how I am doing in Noom
I want to see the ongoing benefits of Noom to my health
I want tailored support while I use Noom
I want to be able to connect my existing health devices to Noom
I do not want Noom to make me feel bad about myself
I want support that is positive and encouraging

As a primary care doctor…
I want to know what digital health tools my patients are using
I want to have easy access to Noom patient data that does not interfere with my current workflows
I want my patients Noom data to be integrated into my existing EHR data
I want to be able to make informed decisions based on my patients Noom data
I want to be able to educate my patients on their health based on Noom
I do not want Noom to negatively impact my workflows

As a dDPP researcher…
I want to be able to identify eligible patients
I want to be able to contact eligible patients using existing technology (eg MyChart)
I want to be able to track our patients Noom use
I want to be able to have detailed patient engagement data from NOOM
I want to remind patients to fill out necessary surveys and onboarding.
I want to know (if) the participants have filled out the relevant study surveys, so I can nudge them if they haven't

B. Human generated output.

Functional Requirements

Retrieve Data Points
- The system shall fetch the following data points from Noom's API: Physical Activity data, including exercise time and steps. Weigh-ins data. Articles read data. Social Activity data, including messages sent to coaches and group posts. The system shall ensure that the fetched data is properly structured and accessible for further processing.

Calculate Compound Metrics
- The system shall combine the "patient daily exercise time" and "daily steps taken" data points into a single metric representing Physical Activity.
- The system shall combine the "messages sent to Noom coaches" and "group posts to Noom's user boards" data points into a single metric representing Social Activity.

Determine Weekly Adherence
- The system shall calculate adherence for each week based on the following thresholds:
- For Physical Activity: Check if the total exercise time for the week is at least 1500 minutes or the total steps taken meet the recommended goal. For weigh-ins: Check if there is at least one weigh-in recorded for the week. For articles read: Check if there is at least one article read during the week. For Social Activity: Check if there is at least one message sent to a coach or one group post made during the week.

Determine Three-Week Adherence
- The system shall analyze the weekly adherence data over the past three weeks to evaluate the overall engagement.
- The system shall calculate the percentage of weeks in which the patient met the adherence criteria for each behavior metric.
- The system shall determine if the patient has met the engagement criteria for each behavior metric over the three-week period.

Generate Nudges
- Based on the calculated adherence for the past three weeks, the system shall categorize the patient's engagement level as "Passing" or "Not Passing."
- The system shall prepare nudge messages tailored to the patient's engagement level over the past three weeks. The nudge messages shall include personalized nudges related to each behavior metric based on the patient's performance.

Schedule Nudge Delivery
- The system shall determine the timing for sending the nudges, following the requirement of sending messages once every three weeks.
- The system shall set up a scheduled task or job to send the nudges to the patient's mobile phone number via SMS.

Monitoring and Evaluation
- The system shall continuously monitor and evaluate the effectiveness of the nudges in promoting patient engagement.
- The system shall collect and analyze data on patient responses, engagement levels, and behavior change outcomes. Based on user feedback and observed outcomes, the system shall iterate on the algorithm and nudge content to improve effectiveness.

Non-Functional Requirements

Performance
- The system shall fetch and process data from Noom's API in a timely manner to ensure real-time nudges generation.
- The nudges delivery task or job shall be scheduled and executed efficiently without delays.

Reliability
- The system shall be robust and resilient, handling potential errors or exceptions during data retrieval, processing, and nudges delivery.
- The nudges delivery mechanism shall have fail-safe measures to ensure message delivery even in case

C. ChatGPT generated output. Prompt: Thanks for this. As a software engineer, give me a well-structured requirements document using the user stories you have already generated in the previous iteration.

The system should be able to connect to the Noom API for seamless data transfer.
The system shall be able to store Noom data into internal data bases
The system should be able to calculate an weekly adherence score using Individual noom patient's data.
The system shall be able to calculate "trends" using 3 weeks of data.
The system shall be able to classify patients using calculated trends.
The system should be able to comment to the twilio API for messaging functionality
The system should be able to send messages to patients based on the adherence score
The system shall be able to automatically trigger the message functionality
The system shall be able to send messages on patient's individual dates such as (3 weeks into the study, birthday, etc)
The system shall be able to send reminders for messages to fill surveys and discuss dDPP study during doctors' visits.
The system shall be able to reply to patient messages
The system shall be able to collect and store message replies in internal databases.
The system should be able to connect to the REDCAP API for secure data transfer and storage
The system shall be able to connect with the EHR (EPIC)
The system shall be able to display the provider UX in EPIC for seamless integration.
The system shall be able to provide an interactive UX for providers to interact with the system.
The UX shall show Noom data, and all doctors requested information

D. Human generated output. Main requirements involve the capacity to query, store, and safely manipulate patients' data; calculate engagement levels; categorize patients' according to engagement level; read, store, and manipulate a behavioral science-based message pool; send personalized messages based on patients' engagement, and generate a reporting dashboard for providers. Requirements include applying high-security data manipulation standards. Data consumed and generated by PAMS needs to be recorded for future evaluation and further analysis of the effect of messages in achieving engagement.

GenAI Solution

We used the output of the "feature construction phase" to inform the GenAI output for requirements. During the initial stages of the prompting phase, we refrained from suggesting solutions, allowing ChatGPT to generate potential solutions autonomously. We reviewed and evaluated these outputs, eliminating impractical or incompatible solution paths that did not align with the intentions or capabilities of our team. Once we reached a satisfactory outcome but faced uncertainty regarding the next steps, we instructed ChatGPT to assume a different "personality" (eg, software architect) and used the previous outputs as a foundation for the new role's initial prompts. Throughout this process, we encouraged each "personality" to seek clarifications by asking questions and provided feedback without biasing toward any predetermined solution. We repeated this process at least four times for each personality type, engaging in a

XSL•FO
RenderX

back-and-forth roleplay with multiple personalities (researcher, architect, and developer), transitioning to a different personality when it became evident that the current one could no longer progress without additional feedback (Figures 4A and 4C).

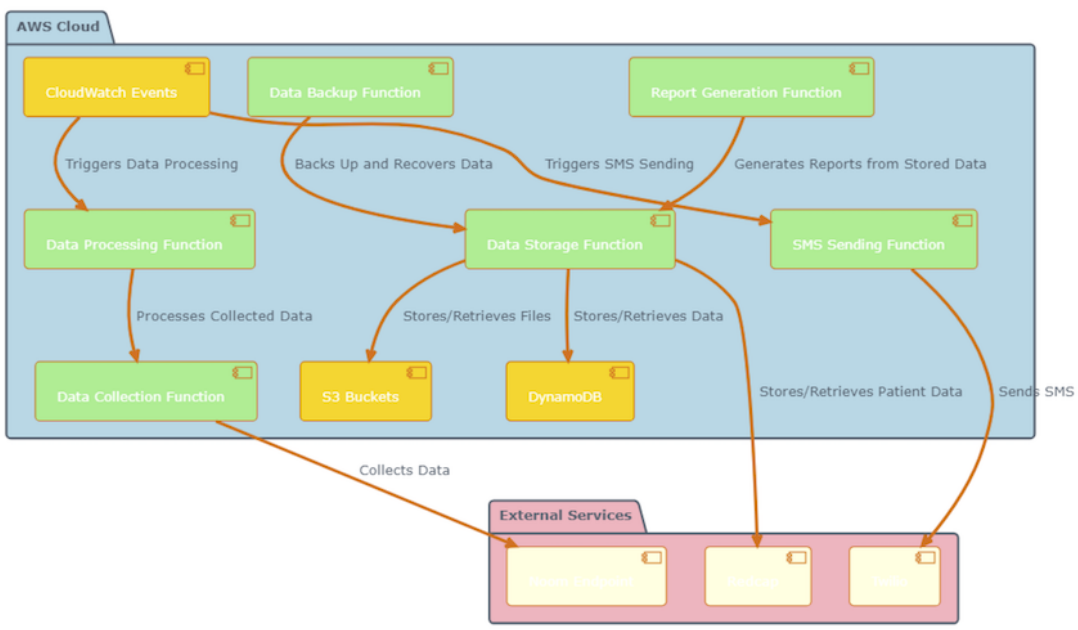
Architectural Design

Human-Generated Solution

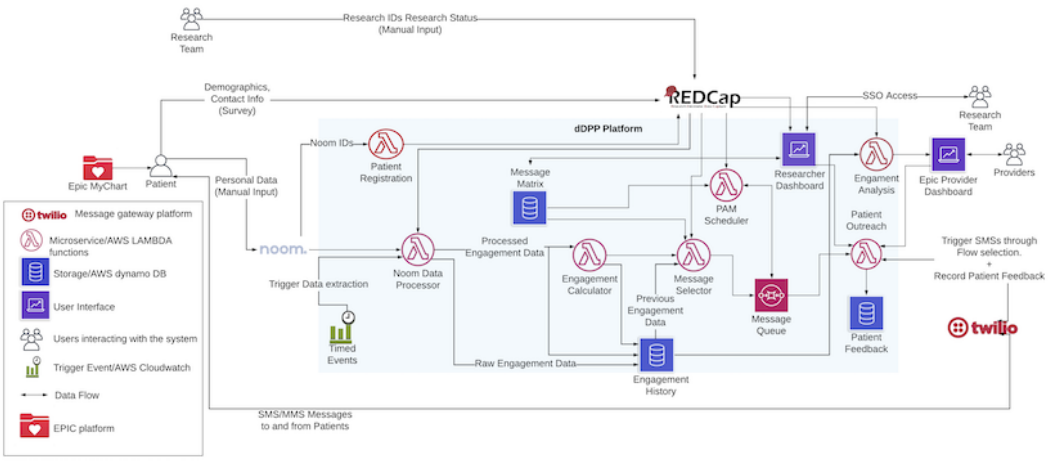
After the requirement phase, our software development team developed the PAMS architectural diagram, which is a graphical representation of the system that includes (1) a set of components (eg, a database and computational modules) that will perform a function required by the system; (2) the set of connectors that will help in coordination, communication, and cooperation between the components; and (3) conditions for how components can be integrated to form the system (Figure 5B).

XSL•FO
RenderX

Figure 5. Architectural design for PAMS. ChatGPT vs human-generated output. AWS: Amazon Web Services; dDPP: digital diabetes prevention program; PAMS: personalized automatic messaging system; REDCap: Research Electronic Data Capture.



A. ChatGPT generated output. Prompt: Thanks for that. With the generated requirements, I will like for you to provide the necessary architectural diagrams for the system. Please consider the following: to develop the system, we want to use AWS and a micro-services approach, also patient data collected from the patients will be stored in RedCap, and we will use Twilio as the SMS gateway to deliver the text messages. Please redo the diagrams using this information.



B. Human generated output. Internal PAMS functionality includes data collection, data analysis, message selection, and message delivery processes, via a modular design where each module is independent and governs a specific activity or functionality within the system. The blue-highlighted area in the figure represents PAMS itself, to include the associated modules, we use symbols. Components outside the blue-highlighted area are external systems and operations that interact with PAMS. These include external applications and research activities that produce data to support PAMS' functionalities or utilize data coming from PAMS. Arrows in the figure show the directions of data flows.

GenAI Solution

For the GenAI-generated architectural design, we leveraged the outputs of the requirement phase and the available ChatGPT plugins to designate the GenAI model as a software engineer and proceeded to develop an architectural diagram. During this process, we engaged in iterative prompting and provided explicit instructions to ChatGPT, specifying the use of Amazon Web Services (AWS) for development, integration of external systems such as Twilio (Twilio Inc) and REDCap (Research Electronic Data Capture; Vanderbilt University), and the adoption of a microservice approach to facilitate the efforts of our development team (Figure 5A).

XSL•FO
RenderX

Code

Human-Generated Solution

PAMS components include several lambda functions that execute its engagement or adherence algorithm, messaging, and data manipulation functionalities. Most of the functions are coded and developed using Python (Python Software Foundation) and Scala (École Polytechnique Fédérale Lausanne) as programming languages. AWS was used for the development of PAMS [12]. Our developers followed our microservice approach design using an event-driven model [13,14]. The main components of PAMS are AWS lambda functions which are triggered by different events such as updates to S3 buckets, modifications on DynamoDB (AWS) tables, or CloudWatch (AWS) events. External interactions of PAMS use application programming interface calls, which secure effective data transfer (Figure 6B).

Figure 6. Code for the function that calculates patient adherence and engagement trends. ChatGPT vs human-generated outputs.

GenAI Solution

To facilitate the generation of the coded solution using ChatGPT, we assigned the role of a software engineer to the model and specifically requested it to generate Scala code for a specific functionality, namely the "calculate engagement trends" function. Consistent with the iterative nature of the GenAI-based software development process, we engaged in a back-and-forth interaction with ChatGPT, iterating over the prompt and its output while providing expert guidance to ensure optimal results. While allowing ChatGPT to generate free text, we evaluated each output for accuracy and adherence to the desired specifications (Figure 6A).

Internal Review of Human Vs GenAI Outputs

The 11 evaluators participated in the output review process. All had familiarity or prior exposure to the original PAMS intervention. Overall, evaluators rated the ChatGPT-produced outputs as positive for the theoretical background and design phase in terms of understandability, usability, novelty, relevance, and efficiency. For these two components, the question about completeness showed the most variability with divided opinion among "agree" and "disagree" and the bias was mostly categorized as "neither agree nor disagree." For the first part of the technical development (user stories and requirement documents), most of the raters found the ChatGPT output positive in terms of understandability, usability, and relevance. In terms of completeness and novelty, requirements were better rated than the user stories which represent an interesting output since requirements are derived from the user stories. We hypothesize that our raters were expecting better user stories, but once these were defined, they considered ChatGPT to be effective at turning these into the requirements. In terms of bias, similar to the theoretical background and design phase, the most popular answer was "neither agree nor disagree." For the more technical pieces of the development that required software engineering knowledge, specifically the architectural diagram and code elements, results showed the highest N/A responses. These higher levels of N/As were associated with lower levels of expertise (eg, coding experience) since only 2 of the 11 evaluators had computer science backgrounds. However, the overall score excluding the N/As was positive for the technical component.

Discussion

Results Summary

This study leveraged ChatGPT-4 to recreate content features and software development of PAMS. ChatGPT served as a usable facilitator for researchers engaging in the software development life cycle, from product conceptualization to

XSL•FO
RenderX

feature identification, and user story development to code generation. GenAI technologies facilitated effective communication and understanding within our multidisciplinary team by providing well-described features and supporting the role of a software engineer. Our findings indicate that the ChatGPT-generated output is comprehensive, albeit with occasional ambiguities that required clarification or adjustment by the research team. The ChatGPT-generated output exhibited a high level of accuracy in capturing the intended requirements. We found that ChatGPT supported a highly efficient development process, producing over 5 days what initially required more than 200 human hours from content and technical experts. The results suggested that by efficiently prompting ChatGPT and leveraging the expertise of our team, we could have significantly reduced the time we invested in initial system modeling and conceptualization phases as well as technical phases of software development (coding). Overall, GenAI technologies like ChatGPT offer a promising approach to efficient software development.

While promising, some significant limitations to ChatGPT's outputs should be noted. In the design phase, while ChatGPT was able to provide general guidance in tool design (eg, app vs web-based vs EHR solution) it was unable to provide evidence to support its rationale for these choices. This lack of reference support has been well-documented and has a material impact on researchers looking to build upon an evidence base for their health technology interventions. Similarly, when asked to provide theoretical frameworks to support behavior change, it offered only a partial list, initially excluding the COM-B (capability, opportunity, motivation, behavior) model upon which the original PAMS intervention was based, and needed prompting from our behavior change expert to provide more specific guidance. In the context of code generation, we focused on testing a specific function, namely the Calculating Patient Engagement feature, which is the core functionality of our software. Initially, we tasked ChatGPT with generating a function to compute a 3-week patient engagement trend. However, the initially generated code deviated from the intended objective and instead calculated a weekly engagement score. Through subsequent iterations, we were able to obtain the desired code. However, the initial attempts exhibited nonidiomatic constructs and contained bugs (no efficient loops and wrong logic). Finally, we observed that ChatGPT overlooked certain suggested features during the design phase, resulting in the generated code occasionally demonstrating unnecessary complexity and disregarding some of the best practices and features of the target programming language. We believe that further iterations would have improved the code quality, encompassing better adherence to coding standards and the inclusion of desired business features, such as handling edge cases and capturing more nuanced engagement trends. Nevertheless, we reached a point of diminishing returns with ChatGPT where we determined that engaging an experienced developer would have expedited the code generation process and ensured a more robust implementation.

These limitations highlight the ongoing importance of human expertise in the development process, especially in scenarios where theoretical expertise, intricate coding practices, and business-specific requirements are involved. The lack of rationale to support the generated results shows the value of having human experts on the team who can interpret the results. ChatGPT needs to be used as a support tool but not the source of truth; thus, we always trusted and relied on human experts to validate the ChatGPT-generated results before moving to the next phase. Overall, it is important to have human experts in the system development process to guide the outputs in terms of reprompting the system (support the decision-making on acceptable output) and ensuring their accuracy. Moreover, results are highly dependent on the quality of the prompts which emphasizes the role of prompt engineering. The results show that well-structured prompts (role + problem description + ask) that infuse human expertise into every iteration are key to obtaining good results (Figure 1). As part of our prompt framework described in the methodology section, results showed that detailed problem explanations, clear asks, and roleplaying are an excellent combination to guide accurate results. We suggest asking ChatGPT questions using different roles, asking for clarification if needed, and in cases of wrong outputs, redirecting the prompts.

Related Work

There is near-universal interest in understanding the impacts of GenAI and large language models (LLMs) on human social structures, including the experience of work and the production of work-related outputs in health care and more broadly [15,16]. In health care, LLMs are poised to impact everything from care delivery experience, diagnostic reasoning and cognitive skills, training and education, and the overall composition of the workforce [17]. These theoretical disruptions are tempered, however, by acknowledging that in its current state, GenAI tools remain suboptimal, with ongoing issues in accuracy, reliability, usability, cost, equity, and ethics.

In commercial spaces, ChatGPT-enabled products designed to assist with coding and software development are already being developed (eg, OpenAI Codex [OpenAI] and CodeGPT [CodeGPT]). These tools can help generate novel code, debug and analyze code issues, assist in code refactoring, and provide code documentation. As yet, however, their usefulness in terms of quality has not been extensively evaluated, and costs and other considerations may make them inaccessible to health care researchers. ChatGPT-enabled tools for front-end design (eg, integrating ChatGPT with Figma [Figma, Inc]), user testing (including synthetic user testing), and prototyping have also been created, all allowing health technology research teams with limited design resources to take advantage of tools from product and experience design to create their interventions. Overall, commercial LLMs have been demonstrated to improve worker efficiency and productivity, through "co-pilot" support services that automate low-skills tasks, organize and present information, and surface insights [18]. Brynjolfsson et al [18] found that a ChatGPT-supported tool providing conversational guidance for customer support agents increased worker productivity by almost 14%. The authors further found that these productivity benefits accrued disproportionately to less-experienced and lower-skill workers, allowing less-skilled or newer workers to experience more rapid gains; the authors posit that high-skill workers may have less to gain from artificial

XSL•FO
RenderX

intelligence assistance due to tacit knowledge reinforcement rather than new knowledge or skill development. Our work suggests that both less-experienced, lower-skill workers and high-skill workers can benefit, with novices benefitting more from new knowledge (if accurate) and skill development and experts benefiting from knowledge validation and offloading of high-effort low-value tasks.

In the academic computer science literature, ChatGPT has been evaluated as a tool for collaborative software design [4], including to improve code quality refactoring, requirements elicitation, and general design solutions [5], and fix programming bugs [19]. Similar findings are reflected in our work, including the caveats of requiring human oversight. Other authors have identified important ethical issues in using GenAI solutions for software engineering, which were not considered in this study [20].

Within health care, a growing body of research has explored the feasibility of GenAI tools (mostly ChatGPT) in a variety of use cases, including answering patient questions [3,21], creating suggestions to optimize clinical decision support [22], generating a history of present illness summaries [23], and overall examination performance [24]. In general, these papers find promising signals for the accurate and acceptable use of GenAI tools, but with many current-state caveats for their optimal, safe, and scaled use. Key areas of concern include reliability (particularly around hallucinations and citation fabrication), reproducibility, and recency of data inputs. While research in this area will continue to grow, as more test cases comparing GenAI performance to that of clinical staff will be undertaken, further work is needed to create validated and generalizable outcome measures. Future work must also ensure that the variety of GenAI tools (including general commercial LLMs, health care–specific LLMs, and internally developed tools) are equally evaluated.

Limitations

There are several limitations to this study. First, no research team members have expertise in prompt generation for GenAI tools; as a result, our prompting reflects the a priori perspectives, biases, and knowledge gaps of our team, and are therefore particularly subject to issues of framing, recall, and confirmation bias that may influence the interpretation of the results. Second, our research team members, who acted as prompt engineers in this study, were highly familiar with the project and participated in the human-based design process; thus, they were aware of what deviations from human-based design to address by reprompting the system. As a result, we have introduced bias in the prompting process and results reflect higher accuracy. Third, the absence of robust tools to objectively measure the "quality" of current ChatGPT outputs poses challenges to accurately and objectively assess its performance. Furthermore, in this case, the output reviewers were not blinded to the human vs ChatGPT outputs, given the complexity of this study and the difficulty in providing enough research context to support independent blind review. Finally, broader limitations of the technology, such as potential hallucinations and concerns about behavioral changes of responses over time, deserve acknowledgment, as they could have implications for the practical applications and long-term viability of GenAI in health care research contexts. Future research efforts should address these limitations to enhance and replicate our findings.

Implications and Future Directions for Exploration

We are considering several future directions for the use of ChatGPT in our digital health intervention development. We envision increasing our expertise in prompt engineering (add expert prompt engineers to the team) to actively use ChatGPT to further develop PAMS features, particularly for additional messaging content. We anticipate this will save our research team considerable time and effort. We may also use ChatGPT to facilitate more time-consuming aspects of our research documentation, including both coding documentation and larger research archival work (eg, meeting minutes and recording intervention decision-making). Overall, we feel ChatGPT and related tools can be effectively leveraged within health care technology research teams with a spectrum of technical expertise, serving to both augment existing skills and supplement skill gaps. For those with expertise in computer science or programming, we imagine ChatGPT can assist by automating high-effort, low-impact tasks or repetitive work that is considered important but often deprioritized as more urgent tasks arise (eg, code documentation). For those without preexisting programming skills, we imagine ChatGPT can offer technical support, including educational tools and skill-building opportunities. Overall, this process will both validate existing knowledge and create new knowledge for teams, as well as potentially improve interteam communication and collaboration.

Conclusions

In this study, we explored the use of the GenAI tool ChatGPT to recreate a novel digital behavior change intervention which our research team had previously developed to support patient engagement and adherence to a commercial dDPP. Specifically, we reviewed and evaluated the capacity and limitations of ChatGPT to support digital health research intervention ideation, design, and software development, finding it a feasible and potential time- and resource-saving tool to support research teams in developing novel digital health products and technologies. At the same time, we identified gaps in ChatGPT outputs that may limit its effective use for both novel and advanced technology developers, particularly around the completeness of outputs. Future directions will include the development of more targeted artificial intelligence–based tools to support health care researchers with all levels of software or engineering skills, as well as the development of improved tools to objectively evaluate GenAI outputs.

XSL•FO
RenderX

Acknowledgments

The National Institute of Diabetes and Digestive and Kidney Diseases (1R18DK118545-01A1, principal investigator: DM) funded this research.

Conflicts of Interest

None declared.

References

1. Risling TL, Risling DE. Advancing nursing participation in user-centred design. J Res Nurs 2020;25(3):226-238. [doi: 10.1177/1744987120913590] [Medline: 34394630]
2. Dave T, Athaluri SA, Singh S. ChatGPT in medicine: an overview of its applications, advantages, limitations, future prospects, and ethical considerations. Front Artif Intell 2023;6:1169595 [FREE Full text] [doi: 10.3389/frai.2023.1169595] [Medline: 37215063]
3. Liu S, Wright AP, Patterson BL, Wanderer JP, Turer RW, Nelson SD, et al. Using AI-generated suggestions from ChatGPT to optimize clinical decision support. J Am Med Inform Assoc 2023;30(7):1237-1245 [FREE Full text] [doi: 10.1093/jamia/ocad072] [Medline: 37087108]
4. Ahmad A, Waseem M, Liang P, Fahmideh M, Aktar MS, Mikkonen T. Towards human-bot collaborative software architecting with ChatGPT. 2023 Presented at: EASE '23: Proceedings of the 27th International Conference on Evaluation and Assessment in Software Engineering; June 14-16, 2023; Oulu, Finland p. 279-285. [doi: 10.1145/3593434.3593468]
5. White J, Hays S, Fu Q, Spencer-Smith J, Schmidt DC. ChatGPT prompt patterns for improving code quality, refactoring, requirements elicitation, and software design. ArXiv Preprint posted online on March 11 2023 [FREE Full text] [doi: 10.48550/arXiv.2303.07839]
6. Rodriguez DV, Lawrence K, Luu S, Yu JL, Feldthouse DM, Gonzalez J, et al. Development of a computer-aided text message platform for user engagement with a digital diabetes prevention program: a case study. J Am Med Inform Assoc 2021;29(1):155-162 [FREE Full text] [doi: 10.1093/jamia/ocab206] [Medline: 34664647]
7. Lawrence K, Rodriguez DV, Feldthouse DM, Shelley D, Yu JL, Belli HM, et al. Effectiveness of an integrated engagement support system to facilitate patient use of digital diabetes prevention programs: protocol for a randomized controlled trial. JMIR Res Protoc 2021;10(2):e26750 [FREE Full text] [doi: 10.2196/26750] [Medline: 33560240]
8. Liu V, Chilton LB. Design guidelines for prompt engineering text-to-image generative models. 2022 Presented at: CHI '22: Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems; April 29-May 5, 2022; New Orleans, LA p. 1-23. [doi: 10.1145/3491102.3501825]
9. Zhou Y, Muresanu AI, Han Z, Paster K, Pitis S, Chan H, et al. Large language models are human-level prompt engineers. ArXiv Preprint posted online on November 03 2022 [FREE Full text]
10. White J, Fu Q, Hays S, Sandborn M, Olea C, Gilbert H, et al. A prompt pattern catalog to enhance prompt engineering with chatGPT. ArXiv Preprint posted online on February 21 2023 [FREE Full text] [doi: 10.48550/arXiv.2302.11382]
11. Wang J, Liu Z, Zhao L, Wu Z, Ma C, Yu S, et al. Review of large vision models and visual prompt engineering. Meta-Radiology 2023;1(3):100047 [FREE Full text] [doi: 10.1016/j.metrad.2023.100047]
12. Nadareishvili I, Mitra R, McLarty M, Amundsen M. Microservice Architecture: Aligning Principles, Practices, and Culture. Sebastopol, CA: O'Reilly Media, Inc; 2016.
13. Alshuqayran N, Ali N, Evans R. A systematic mapping study in microservice architecture. 2016 Presented at: 2016 IEEE 9th International Conference on Service-Oriented Computing and Applications (SOCA); November 4-6, 2016; Macau, China p. 44-51. [doi: 10.1109/soca.2016.15]
14. Mathew S, Varia J. Overview of Amazon Web Services. Amazon Whitepapers. 2014. URL: https://d1.awsstatic.com/whitepapers/aws-overview.pdf [accessed 2024-02-01]
15. Eloundou T, Manning S, Mishkin P, Rock D. GPTs are GPTs: an early look at the labor market impact potential of large language models. ArXiv Preprint posted online on March 17 2023 [FREE Full text] [doi: 10.48550/arXiv.2303.10130]
16. The impact of artificial intelligence on the future of workforces in the European Union and the United States of America. The White House. 2022. URL: https://www.whitehouse.gov/wp-content/uploads/2022/12/TTC-EC-CEA-AI-Report-12052022-1.pdf [accessed 2024-01-18]
17. Thirunavukarasu AJ, Ting DSJ, Elangovan K, Gutierrez L, Tan TF, Ting DSW. Large language models in medicine. Nat Med 2023;29(8):1930-1940. [doi: 10.1038/s41591-023-02448-8] [Medline: 37460753]
18. Brynjolfsson E, Li D, Raymond LR. Generative AI at work. National Bureau of Economic Research. 2023. URL: https://www.nber.org/papers/w31161 [accessed 2024-01-18]
19. Surameery NMS, Shakor MY. Use Chat GPT to solve programming bugs. Int J Inf Technol Comput Eng 2023;3(1):17-22 [FREE Full text] [doi: 10.55529/ijitc.31.17.22]
20. Akbar MA, Khan AA, Liang P. Ethical aspects of ChatGPT in software engineering research. IEEE Trans Artif Intell 2023:1-14. [doi: 10.1109/tai.2023.3318183]

XSL•FO
RenderX

21. Ayers JW, Poliak A, Dredze M, Leas EC, Zhu Z, Kelley JB, et al. Comparing physician and artificial intelligence Chatbot responses to patient questions posted to a public social media forum. JAMA Intern Med 2023;183(6):589-596. [doi: 10.1001/jamainternmed.2023.1838] [Medline: 37115527]
22. Biswas S, Logan NS, Davies LN, Sheppard AL, Wolffsohn JS. Assessing the utility of ChatGPT as an artificial intelligence-based large language model for information to answer questions on myopia. Ophthalmic Physiol Opt 2023;43(6):1562-1570 [FREE Full text] [doi: 10.1111/opo.13207] [Medline: 37476960]
23. Nayak A, Alkaitis MS, Nayak K, Nikolov M, Weinfurt KP, Schulman K. Comparison of history of present illness summaries generated by a Chatbot and senior internal medicine residents. JAMA Intern Med 2023;183(9):1026-1027. [doi: 10.1001/jamainternmed.2023.2561] [Medline: 37459091]
24. Strong E, DiGiammarino A, Weng Y, Kumar A, Hosamani P, Hom J, et al. Chatbot vs medical student performance on free-response clinical reasoning examinations. JAMA Intern Med 2023;183(9):1028-1030. [doi: 10.1001/jamainternmed.2023.2909] [Medline: 37459090]

Abbreviations

AWS: Amazon Web Services
COM-B: capability, opportunity, motivation, behavior
dDPP: digital diabetes prevention program
GenAI: generative artificial intelligence
LLM: large language model
N/A: not applicable
PAMS: personalized automatic messaging system
REDCap: Research Electronic Data Capture

Edited by A Kushniruk; submitted 18.09.23; peer-reviewed by S Mulvaney, I Oakley-Girvan; comments to author 20.11.23; revised version received 27.11.23; accepted 15.12.23; published 06.03.24.

Please cite as:
Rodriguez DV, Lawrence K, Gonzalez J, Brandfield-Harvey B, Xu L, Tasneem S, Levine DL, Mann D
Leveraging Generative AI Tools to Support the Development of Digital Solutions in Health Care Research: Case Study
JMIR Hum Factors 2024;11:e52885
URL: https://humanfactors.jmir.org/2024/1/e52885
doi:10.2196/52885
PMID:38446539

©Danissa V Rodriguez, Katharine Lawrence, Javier Gonzalez, Beatrix Brandfield-Harvey, Lynn Xu, Sumaiya Tasneem, Defne L Levine, Devin Mann. Originally published in JMIR Human Factors (https://humanfactors.jmir.org), 06.03.2024. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on https://humanfactors.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Testing Two Online Symptom Checkers With Vulnerable Groups: Usability Study to Improve Cognitive Accessibility of eHealth Services

Kaisa Savolainen[1], DA (Art and Des), MSc (tech); Sari Kujala[1], PhD

Department of Computer Science, Aalto University School of Science, Espoo, Finland

Corresponding Author:
Kaisa Savolainen, DA (Art and Des), MSc (tech)
Department of Computer Science
Aalto University School of Science
Konemiehentie 2
Espoo, 02150
Finland
Phone: 358 503101457
Email: kaisa.savolainen@aalto.fi

Abstract

Background: The popularity of eHealth services has surged significantly, underscoring the importance of ensuring their usability and accessibility for users with diverse needs, characteristics, and capabilities. These services can pose cognitive demands, especially for individuals who are unwell, fatigued, or experiencing distress. Additionally, numerous potentially vulnerable groups, including older adults, are susceptible to digital exclusion and may encounter cognitive limitations related to perception, attention, memory, and language comprehension. Regrettably, many studies overlook the preferences and needs of user groups likely to encounter challenges associated with these cognitive aspects.

Objective: This study primarily aims to gain a deeper understanding of cognitive accessibility in the practical context of eHealth services. Additionally, we aimed to identify the specific challenges that vulnerable groups encounter when using eHealth services and determine key considerations for testing these services with such groups.

Methods: As a case study of eHealth services, we conducted qualitative usability testing on 2 online symptom checkers used in Finnish public primary care. A total of 13 participants from 3 distinct groups participated in the study: older adults, individuals with mild intellectual disabilities, and nonnative Finnish speakers. The primary research methods used were the thinking-aloud method, questionnaires, and semistructured interviews.

Results: We found that potentially vulnerable groups encountered numerous issues with the tested services, with similar problems observed across all 3 groups. Specifically, clarity and the use of terminology posed significant challenges. The services overwhelmed users with excessive information and choices, while the terminology consisted of numerous complex medical terms that were difficult to understand. When conducting tests with vulnerable groups, it is crucial to carefully plan the sessions to avoid being overly lengthy, as these users often require more time to complete tasks. Additionally, testing with vulnerable groups proved to be quite efficient, with results likely to benefit a wider audience as well.

Conclusions: Based on the findings of this study, it is evident that older adults, individuals with mild intellectual disability, and nonnative speakers may encounter cognitive challenges when using eHealth services, which can impede or slow down their use and make the services more difficult to navigate. In the worst-case scenario, these challenges may lead to errors in using the services. We recommend expanding the scope of testing to include a broader range of eHealth services with vulnerable groups, incorporating users with diverse characteristics and capabilities who are likely to encounter difficulties in cognitive accessibility.

(JMIR Hum Factors 2024;11:e45275) doi:10.2196/45275

KEYWORDS

eHealth; online symptom checkers; usability; cognitive accessibility; web accessibility; qualitative research

XSL•FO
RenderX

Introduction

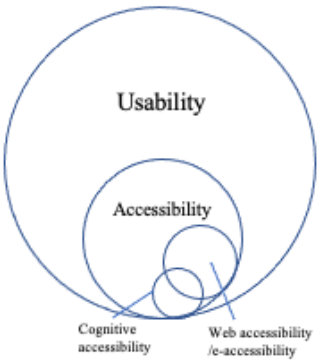
Background

Given the widespread use and popularity of eHealth services, there is a growing need for more accessible services to all potential user groups [1]. In recent years, more emphasis has been placed on accessibility and inclusion; for example, the European Union Accessibility Act has been incorporated into and enforced as national law since June 2022 [2]. As health care services are often public services, it is important that they serve a broad range of users. Furthermore, usability has been recognized as a key component of eHealth applications, and users may face problems with using the applications due to their health conditions [3]. In addition, patients with chronic illness have been reported to encounter more cognitive challenges [4]. Thus, extra attention should be paid to the usability of eHealth applications.

Universal design and design for all address these requirements by aiming at designing services that are usable by and accessible to all user groups regardless of their age, abilities, or possible disabilities [5]. Usability is a high-level term that indicates how a system can be used by specified users in a certain context of use to achieve specific goals with regard to effectiveness, efficiency, and satisfaction [6]. Accessibility, which is a part of usability, describes how a system can be used by people with the widest range of needs, characteristics, and capabilities [6,7]. Thus, accessibility covers all sorts of users with different limitations. A concept that has been addressed by several research papers [8,9] is web accessibility (or e-accessibility), which refers to the accessibility of web services.

In this paper, we address cognitive accessibility, which refers to accessibility beyond physical and sensory capabilities, and thus takes into account varied human characteristics such as intellectual disabilities, attention difficulties, reading problems, autism spectrum disorders, and low language skills [10]. Cognitive accessibility is an important aspect of web accessibility as it involves a large number of users and has a high impact on usability [10]. A summary of the relationship between these concepts is presented in Figure 1.

Figure 1. The relation of cognitive accessibility to usability and accessibility. Note that the sizes and positions of the circles are indicative.



This research focuses on cognitive accessibility within the context of the 2 most frequently used online symptom checkers in Finnish public primary care across numerous municipalities in Finland. Online symptom checkers are used by people seeking health-related guidance, and these services typically provide an urgent assessment and suggest guidance based on the symptoms reported by the user [11]. Patients can use the 2 examined symptom checkers to book appointment times for doctors and laboratory tests or obtain medical help for the most common health issues. First, patients report their symptoms and submit them to the health care center through the symptom checker. Health care professionals receive patient inquiries with an urgency rating, decide on actions to be taken, and inform patients.

Patients are generally highly satisfied with symptom checkers, but younger and more highly educated people have been more likely to use them [11]. For example, symptom checkers enable patients to access health care anytime and anywhere. Therefore, it is essential to ensure that all user groups, including individuals in vulnerable situations, can use these services effectively. Symptom checkers can also empower users as a means of facilitating their health care [12]. However, the accuracy of the symptom checkers depends on how well patients are able to communicate their symptoms when using the tools [13]. As these services spread and are used by a wider range of individuals, it is crucial to also evaluate their usability and accessibility with a more diverse set of users.

Prior Work

Vulnerable Groups

Many public eHealth services and their poor usability and accessibility can cause challenges for certain user groups [14]. These user groups are, thus, in a potentially vulnerable situation in using the service and at risk of digital exclusion [1]. This is especially problematic because research has shown that digital exclusion can cause social exclusion [15]. Public health services must, thus, address the needs of potentially vulnerable groups, including people who are disadvantaged by health, economic, cultural, or social conditions [16], such as older adults, migrants, mental health service users, and the unemployed [16,17].

Older adults are the largest group to face challenges in using digital health services [18,19]. As people age, their cognitive abilities may weaken, with cognitive load being identified as the most significant accessibility barrier for older adults [20]. Memory changes can also affect learning, information processing, and language comprehension [21,22]. Additionally, older adults often struggle with focusing their attention, particularly when multitasking [21,22]. Moreover, older age

XSL•FO
RenderX

groups tend to use eHealth services less frequently than younger demographics. A Finnish study examining an online symptom checker (referred to as service A in this study) observed that individuals aged 20-39 years used the service more actively compared with older age groups, relative to their representation in the population [23]. This suggests that enhancing service usage entails prioritizing usability and accessibility from the perspective of older users as well.

Migrants represent a growing demographic that often faces challenges when accessing health services in their new country of residence [1]. Language barriers and a lack of digital skills are common issues encountered by this group [1]. Additionally, individuals with intellectual disability are another vulnerable population impacted by the digitalization of health services [24]. They have been noted to experience more difficulties in finding information on the internet and understanding online information compared with the general population [25].

Previous research suggests that vulnerable groups, such as older adults and individuals with mild intellectual disabilities, encounter cognitive challenges when using technology [26,27]. Therefore, the development of more accessible eHealth services would enable these groups to access health information more easily [25,28], thereby enhancing their sense of empowerment concerning their health issues.

The preferences or needs of older adults or individuals with mild intellectual disabilities are often overlooked in the majority of eHealth studies [29,30]. It is imperative to better consider these user groups during the design of eHealth services [17,28]. Many eHealth applications could greatly benefit from the application of universal design principles [29], which facilitate understanding the needs of potentially vulnerable groups and inform the design of more inclusive and usable services [31,32]. Consequently, this enables vulnerable groups to derive as much benefit from eHealth systems as the rest of the population [33,34]. Indeed, universal access approaches can offer benefits to anyone [35]. Therefore, to gain a better understanding of the challenges faced by vulnerable groups when using services, it is essential to conduct testing with a diverse group of users.

Usability Testing of Symptom Checkers

The usability of symptom checkers has been examined in prior research; however, there has been limited emphasis on potentially vulnerable user groups, such as older adults, migrants, and those with intellectual disability [36-38]. Moreover, research on usability in the eHealth domain frequently concentrates on quantitative aspects (eg, the number of errors, task completion times, and usability questionnaires) and typically involves a large number of users [12,36,38]. However, the qualitative aspect of usability studies is also crucial for gaining a deeper understanding of the thoughts and reasons behind errors, as well as capturing the patient's perspective at a broader level [3,39]. Additionally, while a System Usability Scale (SUS) questionnaire provides a numeric score for experienced usability, it alone is not adequate for evaluating usability. Instead, it should be complemented with other measures, such as task completion rates or more qualitative approaches, to ascertain which aspects of a service require improvement and how best to address them [39].

Marco-Ruiz et al [13] conducted research on symptom checkers and emphasized the significance of testing with real users to comprehend the cognitive processes involved when using a new system to record health data. Furthermore, they noted that the user base accessing symptom checkers is highly diverse, with some individuals possessing higher health literacy and experience in recording online information, while others may have very limited or no experience [13].

Goal of the Study

The goal of our study is to gain a deeper understanding of cognitive accessibility in the context of eHealth services. Therefore, our paper focuses on addressing the following research questions:

- What kind of challenges do vulnerable groups face in using eHealth services?
- What needs to be considered when testing with vulnerable groups?

The structure of this paper is as follows: In the next section, we describe the methods used in this study, followed by the presentation of results. Subsequently, we discuss the findings and overarching contributions of this study, concluding with our final remarks.

Methods

Approach and Researcher Background

Our qualitative study adopts a case study approach, wherein the cognitive accessibility of eHealth services was assessed through usability testing of 2 online symptom checkers. The research team comprised 3 researchers: The first researcher, a human-computer interaction student, conducted the initial 8 tests as part of their master's thesis work. Subsequently, a second researcher, a senior researcher with expertise in human-computer interaction (who served as the thesis advisor), conducted the remaining 5 tests. Additionally, a third senior researcher with backgrounds in human-computer interaction and eHealth oversaw the entire study.

Context and Study Setting

We conducted a usability test of 2 Finnish online symptom checkers in 2 phases in Finland during the Spring and Fall of 2021. The tested services were Omaolo (DigiFinland Oy) [40] and Klinik Access (Klinik Healthcare Solutions Oy) [41], which are the 2 most-used symptom checkers in Finnish public primary care. Omaolo has been actively used since 2019, while Klinik Access, which is also used internationally, has been in use since 2015. Both services are designed to assist patients in obtaining appropriate care. Users answer a set of questions regarding their symptoms, following which the symptom checkers use artificial intelligence to assess the urgency of care. If necessary, the services guide patients to contact emergency care services.

The Omaolo symptom checker comprises 15 specialized symptom checkers tailored for different types of symptoms, along with a generic symptom checker. Each symptom checker prompts the user with a specific set of questions and subsequently recommends the next steps they should take. Additionally, if the user provides their home municipality, the

XSL•FO
RenderX

service displays recommended actions specific to the area, offers contact details, and may even facilitate direct contact with health care professionals if deemed necessary. The Omaolo symptom checker served as the primary COVID-19 symptom checker in Finland, enabling users to schedule appointments for COVID-19 tests. Consequently, its user base experienced a significant surge [23].

The Klinik Access symptom checker enables users to initially select the part of the body where their main symptoms are located. Subsequently, it prompts for more specific symptoms. The responses can then be forwarded to the medical staff responsible for the patient's care before their appointment, ensuring the patient is directed to the appropriate type of health care professional. The primary distinction between these services lies in their user interface (UI): Klinik Access features a more visual UI with a list of clickable symptoms, whereas Omaolo presents users with multiple-choice questions describing the symptoms. Henceforth, the Omaolo service will be denoted as service A, and Klinik Access will be referred to as service B. It is important to note that both services are classified as medical devices and must adhere to specific safety requirements, such as repetitive questions, which may impact usability.

Sampling Strategy

Purposive sampling [42] was used to recruit participants, who were sourced through personal contacts and various associations representing the targeted user groups. These associations included initiatives such as the Selkeästi meille, which focuses on enhancing cognitive accessibility, and Väylä ry, which is dedicated to improving the employment opportunities of individuals with intellectual disability. It is important to note that the test facilitator did not have a close personal relationship with the participants, such as being a friend or family member, during any of the test sessions.

A total of 13 participants were recruited to partake in the study. Notably, an evaluation of sample sizes within the field of human-computer interaction has indicated that 12 is the most common sample size for usability studies [43].

Ethical Approval

The study received approval from the ethical review board of Aalto University (D/902/03.04/2021). Each participant provided informed consent by signing a consent form after confirming their understanding of the study's purpose and how their information would be handled. Reporting has been conducted in such a manner that individual participants cannot be identified.

Data Collection Methods

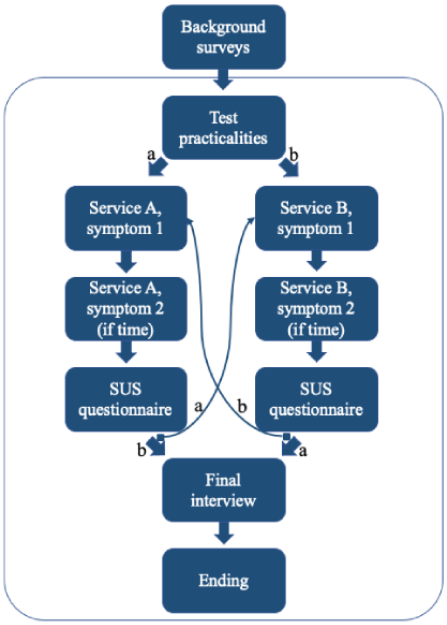
Overview

The main methods used in this study were thinking aloud, observations, questionnaires, and semistructured interviews. Before the actual tests, a pilot test was conducted to identify any potential inconsistencies and to ensure that the questions and instructions were comprehensible. Minor adjustments to the test setup were made based on the findings from the pilot test.

Test Procedure

An overview of the test sessions is presented in Figure 2.

Figure 2. An overview of the usability test sessions with older adults, mildly intellectually disabled individuals, and nonnative speakers (N=13). Half of the participants started with Service A and the other half with Service B.



Each participant tested both services, and the order of service usage was counterbalanced. During the testing phase, participants were presented with 2 symptom vignettes, each providing a brief description of the symptoms they were instructed to imagine having. These vignettes were used 1 at a time. Participants were then asked to open the service and imagine they had the symptoms described in the first vignette, aiming to determine how they should proceed. The vignettes and mode of distribution between the participants are presented in Multimedia Appendix 1.

XSL•FO
RenderX

After using the first service, participants were instructed to take the second vignette and attempt to use the service again. However, if the first part of the test had exceeded 40 minutes, the second vignette was omitted for the first service to prevent the overall test time from exceeding 90 minutes. Following their interaction with each service, participants were asked to evaluate the respective service.

After testing the first service, participants were instructed to open the second service and follow the same procedure. Upon completion of both testing phases, participants were asked to compare the 2 systems and select the one they preferred.

Data Collection Instruments

Test Sessions

The test sessions were conducted via the Microsoft Teams videoconferencing platform, which facilitates screen sharing, screen recording, and voice recording functionalities. The decision to conduct remote testing was primarily influenced by the COVID-19 pandemic situation, but it also aligned well with the nature of the tests, as the services being evaluated were online. Participants used their personal computers to access the services during the testing sessions.

Symptom Vignettes

To streamline the usability test and eliminate the necessity for participants to input their personal medical information into the services, each participant was provided with 2 standardized clinical vignettes featuring predefined symptoms. These vignettes were selected from a list compiled by Semigran et al [44], encompassing a total of 6 conditions with varying severity levels. The selection included conditions with different severity levels to account for the fact that individuals may use symptom checkers in both urgent and nonurgent situations [45].

In line with the recommendations provided by Semigran et al [44], the selected vignettes encompassed 3 categories of triage urgency: conditions necessitating emergency care, conditions warranting nonemergency care, and conditions deemed unnecessary for medical visits, thus manageable with self-care. Moreover, we opted for conditions commonly observed within the age group under study to ensure relevance. These conditions encompass ailments such as acute bronchitis, back pain, and meningitis. To ensure clarity and relevance to the participants, the selected conditions were translated from English to Finnish and simplified. The English versions of the vignettes used can be found in Multimedia Appendix 2.

Background Questionnaires

Before the actual test session, participants were requested to complete a brief background survey and the health literacy survey HLS-EU-Q16 [46]. The background information collected were the participant's gender; age; the frequency of doctor visits in the preceding 2 years; the number of doctor-diagnosed medical conditions; their previous usage frequency of digital health care services; and their frequency of digital device usage, such as smartphones or computers. These questions aimed to ascertain whether participants met the study's target demographic criteria in terms of age and their ability to independently use electronic devices such as computers. The health literacy survey provided insights into participants' understanding of health-related topics.

Interview and Questionnaire

After interacting with each service, participants were asked to evaluate the tested services. This involved administering an SUS questionnaire [47] to gauge the perceived usability of the system, as well as posing 4 interview questions:

- Would you use the service again in the future?
- Were the summary and the instructions about what to do next clear enough?
- Would you actually follow the instructions given?
- Given the option, would you use the service using your phone?

Data Processing and Analysis

The test sessions were recorded using Microsoft Teams. The voice recordings of the initial 8 tests were transcribed in full, while for the remaining 5 tests, notes were taken from the recordings, and user comments were documented to streamline the process. An experienced researcher could identify the issues encountered by users as well as their comments without requiring a complete transcription. The notes and transcriptions underwent anonymization. Qualitative content analysis was used in this study. Using the notes and recordings, all usability issues were identified and compiled. This encompassed problems mentioned by participants as well as those observed during testing or evident from the recordings. The identified usability problems were coded and categorized based on their similarities. When new problems were identified, they were compared with existing ones, and if deemed similar, they were grouped under the same code. Eventually, these groups were consolidated under higher-level descriptive categories. Furthermore, user comments were collected to bolster the analysis and reporting process.

The background questionnaires were analyzed by aggregating the responses to obtain an overview of participant characteristics. Additionally, the health literacy surveys were analyzed according to the guidelines [46] to determine the groups to which participants belonged. The SUS questionnaires were analyzed by computing the SUS scores as per the guidelines [47], resulting in scores of up to 100 points, which were then compared with the general score.

To ensure the quality and trustworthiness of the study, a senior researcher (the second author) supervised the entire research process and provided support for the analysis work. Two other researchers (the first author and the master's thesis worker) conducted the actual tests and analyzed the data. Therefore, a total of 3 researchers participated in the process, ensuring that data gathering and processing proceeded appropriately.

Results

Overview

The subsequent sections present the principal findings of the study. We commence with an overview of the participants' characteristics, followed by an examination of the identified

XSL•FO
RenderX

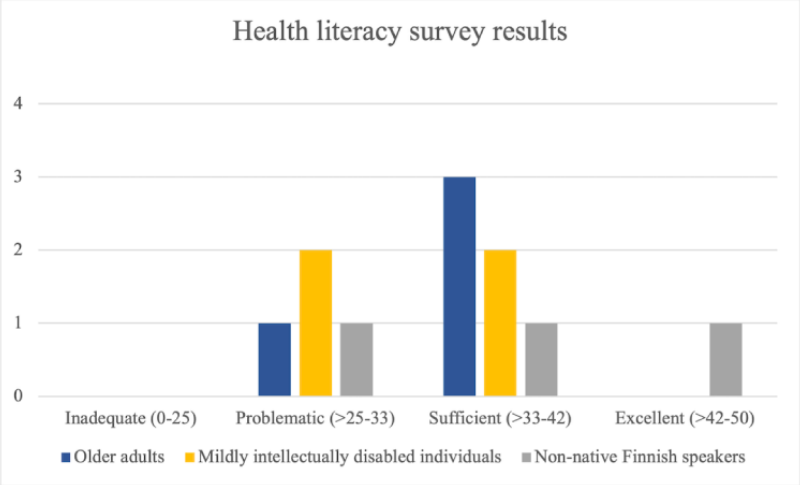
usability issues. Finally, we present additional findings that emphasize the characteristics of these user groups.

Test Participants

A total of 13 individuals participated in the study in Finland. Among them, 4 were individuals with mild intellectual disability, 4 were older adults (aged 75-79 years), and 5 were nonnative Finnish speakers. Therefore, all test users potentially encountered cognitive accessibility challenges with the services. The background characteristics of the participants are detailed in Multimedia Appendix 3.

The HLS-EU-Q16 questionnaire results were calculated in accordance with the guidelines [46], with each participant receiving a score corresponding to a quartile representing their health literacy level. The results were computed only for participants who responded to at least 80% of the questions, as recommended in the guidelines [46]. The questionnaire includes an "I don't know" answer option, which was interpreted as the question not being answered. Consequently, the results of 2 of the nonnative participants were excluded, as they chose this answering option too frequently. Figure 3 depicts the distribution of health literacy among the 3 groups.

Figure 3. The results of the HLS-EU-Q16 health literacy assessment of older people, mildly intellectually disabled individuals, and non-native speakers divided into four categories.



Usability Problems

Cognitive Accessibility Issues

The study identified a total of 65 usability problems with the 2 systems. Specifically, 36 usability problems were discovered with service A, while 29 problems were identified with service B. These issues occurred across 99 and 91 individual user instances, respectively. The problems were classified into 14 usability problem categories. A comprehensive list of the usability problem categories is provided in Multimedia Appendix 4. For the purpose of this discussion, we will focus on issues related to cognitive accessibility, primarily concerning terminology, text volume, and UI clarity.

Terminology-Related Issues

The most prevalent issues were associated with terminology and answering options. eHealth services frequently incorporate specialized language and specific terminology, posing challenges for users with cognitive limitations. Nearly all users encountered confusion with certain terms or inadvertently mixed them up with similar ones. Furthermore, lengthy words and extensive blocks of text, such as lengthy paragraphs, presented challenges, a sentiment that was also echoed during the interviews. Users with cognitive restrictions often encounter challenges when confronted with long words and extensive passages of text.

As one user commented,

> *It takes time to go through all the texts.* [ID10, nonnative]

Related issues were reported and commented on by users across all user groups. In addition to contributing to usability problems, these issues slowed down the usage of the services and occasionally led users to select incorrect symptoms.

Issues Related to the Clarity of the UI

Another area where users encountered difficulties was with the visibility of information and the lack of clarity in the UI. It is crucial for the most important information and elements of the UI to be clearly visible, facilitating easy comprehension for users. Additionally, problems arose when users' attention was diverted to unimportant features. These issues are especially pronounced among user groups with cognitive difficulties, as they require additional attention to comprehend the content and must focus more intently. Furthermore, some users found the input methods challenging; initially, they struggled to discern the type of information required for input in a field and how the inputting should be performed.

The most prevalent individual usability problems we identified regarding the logic and functionality of the UI, observed across all 3 user groups, are detailed in Textbox 1.

XSL•FO
RenderX

Textbox 1. Individual usability problems identified.

1. **Users making an incorrect selection due to an item being highlighted in the user interface:**

   Service A was, at the time of the study, the prevalent symptom checker for COVID-19 in Finland; COVID-19 was highlighted at the top of the home page of service A and was thus the first item to attract the users' attention and be selected.

2. **Difficulties in making the correct selection from a long list of items:**

   Service A had a list of 15 symptom checkers from which the user had to choose, making it difficult for the users to select the correct symptom checker to continue with.

3. **Not remembering what questions needed to be answered after the questions disappeared:**

   Service B presented questions as placeholders to describe symptoms in open answers, and these questions disappeared when the user started typing in the field; as a result, the user might not fully describe their symptoms.

4. **Being confused by long lists of apparently uncategorized symptoms:**

   Service B had long lists of symptoms as selectable buttons that seemed to be unorganized and caused anxiety and confusion.

5. **The logic and functionality of submenus were not understood by the users**:

   Service B had additional submenus and dialog boxes that were not fully understood by the users. There was a small arrow that opened the submenu and the logic of how the items were selected or the submenus opened was unclear.

6. **The users did not understand the logic of the input fields that combined several user interface items:**

   The way in which service B required the duration of symptoms to be input meant that the user needed to enter the number in one field and then select the unit from different options. However, the unit selection was not clearly related to the textbox where the user inputs the number.

These individual usability problems highlight issues with how information is presented to users, with clarity being particularly emphasized among this user group. In some instances, the selection or input options were unclear, and the services featured lengthy lists of symptoms.

Clarity was a recurring theme in several test sessions. As one user commented:

> *...if you think about this in real life, if you have a fever and you're doing this and you start to scroll all these selection choices and you're evaluating which one would fit best, the options are quite broad, so it might be quite difficult to do in practice...* [ID7, user with mild intellectual disability]

Similarly, one user suggested:

> *I don't know you could kind of put those in order like one row and another row, these are quite...your eye kind of jumps, but otherwise those are clear.* [ID2, older adult]

One user preferred the structure of service A and, again, referred to the clarity with which the information is presented:

> *Well, [I prefer Service A] because it was maybe better organized, there was one thing and one question and then one answer. After this, the next question and so on. In the other one [Service B], you had to read all the small boxes and look for your symptom. [...]* [ID8, user with mild intellectual disability]

> *Well, maybe what is the most [difficult], this one had so many small boxes that at least for me, it was difficult to find my own symptom, the one I needed to select from there. So, if I wanted to know what fit me, I had to read through them all and then, since they are not in any order, they just are there, I had to read them all, to see if I could find the one I have at the moment.* [ID8, user with mild intellectual disability]

It is worth noting that the symptoms were arranged in alphabetical order; however, the layout was such that users did not realize this ordering method had been used.

Differences Between the User Groups

Some differences between the user groups were evident, although the majority of the usability problems were consistent across all user groups. Nonnative Finnish speakers found the service to be particularly slow to use, often taking an extended period to read the texts. One user commented regarding service B that:

> *Reading and writing text is not easy for an immigrant. When you can click on an item it is easy, you don't have to write.* [ID13, nonnative]

The older adults did not encounter as many issues with longer texts. Instead, they faced more challenges in understanding the logic of the services and remembering to scroll down to view all the provided information. However, this scrolling also frustrated some nonnative users; as one user commented:

> *And again, we're scrolling, this is terrible!* [ID12, nonnative]

The task completion times were also measured and presented for the initial tasks of both services. As depicted in Table 1, aside from the older adults, there were no significant differences in the completion times between the services. However, for the older adults, service B, which featured more clickable elements to choose from, appeared to be quicker to use. Nonnative speakers took the longest time to complete the tasks, primarily because they often needed to translate some of the terms used in the services. Three of the users used an online translator (eg, Google Translator), and at times, users asked the facilitator about specific terms. Overall, the task completion times were

XSL•FO
RenderX

quite lengthy, suggesting that these user groups require ample time to use these services effectively.

Table 1. Average task completion times (first task) for both services. For older adults there was a clear difference in favor of service B; for the 2 other groups service A got a slightly better time.

Task completion times	Service A, hh:mm:ss[a]	Service B, hh:mm:ss
Older adults	0:14:30	0:08:00
User with mild intellectual disability	0:14:54	0:16:00
Nonnatives	0:15:46	0:18:16

[a]hh:mm:ss: hours:minutes:seconds.

For the few users who had the opportunity to test the services twice, the second time was generally much faster than the first, indicating good learnability. As one user mentioned:

> *Now I know that I need to select this and not the other, which I didn't know previously.* [ID8, user with mild intellectual disability]

The SUS scores are provided in Multimedia Appendix 5, illustrating how participants evaluated the usability of the services. The SUS score ranges from 0 to 100 points. It has been assessed for numerous services, and according to Bangor et al [48], a satisfactory SUS score is above 70, with superior products typically scoring 80 or higher. However, it is important to note that the interpretation of SUS scores can vary depending on the type of product and its development phase. When evaluating the SUS scores of the tested services, which are predominantly below 75, it is evident that the perceived usability was not considered very good, except for nonnative Finnish speakers, as their scores hovered around 80.

From the interviews, we found that older adults tended to prefer computers over mobile devices when using the symptom checkers, whereas nonnative speakers mostly preferred mobile devices. The preference among users with mild intellectual disability was evenly divided. Nonetheless, the advantage of this type of online symptom checker was evident, as all participants expressed willingness to use the services again. The nonnative participants particularly valued a service that enabled them to input information at their own pace, as opposed to speaking on the phone. However, their preference for the service they would use was fairly evenly split, with no clear consensus: 7 participants favored service A, while 6 participants favored service B.

Discussion

Principal Findings

Testing for cognitive accessibility with 2 symptom checkers revealed that older adults, individuals with mild intellectual disability, and nonnative speakers may encounter numerous challenges when using the services. Primarily, problems arise concerning the terminology used. This highlights the need for greater emphasis on ensuring that the vocabulary used in the health sector, while specialized, remains understandable to a broad audience when services are intended for universal use. Furthermore, complications arose from the intricate structure and layout of the services. The significance of simplifying services, minimizing lengthy lists, and using more understandable terminology was highlighted in nearly all the test sessions. Implementing these improvements to the services would likely benefit a broader range of users [5].

There were distinct differences observed among the 3 user groups. Primarily, nonnative speakers assigned notably higher usability ratings to the services compared with the other 2 groups. One possible reason for this could be their overall satisfaction with the existence of such services, which enable them to seek help for their health issues without having to converse over the phone in a language that is not their native tongue.

One notable distinction between the user groups pertained to their preference for using either a computer or a mobile device. It was evident that older adults favored using computers, likely because of their larger screens and the familiarity that older adults have with them. Conversely, most nonnative Finnish speakers showed a preference for mobile devices, with some noting that they solely rely on their mobile devices and do not even own a computer. This preference may be influenced in part by financial constraints, which limit the number of devices a person can afford. Additionally, in our sample, older adults encountered fewer difficulties with processing long pieces of text compared with the other groups.

The promotion of online symptom checkers as a means to decrease unnecessary clinic visits [13] underscores the importance of ensuring they do not inadvertently increase contact with health care staff. Therefore, greater attention should be directed toward enhancing the cognitive accessibility of these tools, thereby enabling a wider range of users to use them effectively. In this study, users' incomplete understanding of the questions or answer options led them to select additional symptoms, resulting in more serious care recommendations and advising users to seek emergency health care.

In ensuring the cognitive accessibility of eHealth services, it is imperative to involve vulnerable groups in testing. Testing with vulnerable groups provides valuable insights. First, it emphasizes the need for well-planned test sessions with a manageable number of tasks. This approach ensures that participants can fully engage and provide meaningful feedback without being overwhelmed. All of these groups required considerable time to complete the test tasks, with most participants unable to finish both planned tasks with either service. Moreover, they necessitated more detailed instructions and support during the test sessions, as many participants within these groups were not at ease with using eHealth services.

XSL•FO
RenderX

Based on the findings of this study and as supported by the broader universal design literature [5], several design guidelines can be outlined. Foremost among these is the emphasis on clarity. (1) The options provided to the user should be clear and understandable. The user should understand what the differences between different options are and what actions are available for them. (2) It should be made clear to the user where they should be focusing on. This is particularly important in services that contain a lot of information and options. (3) Long or uncommon words and difficult compound words should be avoided. This is especially relevant in health-related terminology, as the user might not understand the special terms and might confuse different terms. (4) Navigating the services should be easy and effortless. The user should be presented with as few options as possible, and excessive scrolling should be minimized. This is because the user may inadvertently overlook relevant information.

Limitations

There are, naturally, some limitations to this study. First, the sample size of 13 participants was rather small, albeit quite typical for this type of qualitative study [43]. However, given the diverse nature of the user group and potential challenges related to cognitive accessibility, a more diverse participant pool could have been beneficial. Specifically, a wider age range of older adults could have been tested, considering their versatility as a group. Additionally, nonnative Finnish speakers could have been recruited from a more geographically diverse range of countries of origin. Moreover, testing should involve other diverse human characteristics, such as neurodiversity (including conditions such as attention-deficit/hyperactivity disorder, attention-deficit disorder, and various forms of autism). Given society's rapid transition toward digitalized services, it is crucial to broaden the scope to include other groups at risk of digital exclusion.

Another limitation of this study is its focus on only 2 online symptom checkers. While the range of available online symptom checkers is already extensive, it is important to include testing of other eHealth services designed for use by all citizens. Additionally, this study only examines a limited list of symptoms and assesses usage on a 1-time or 2-time basis.

In conclusion, we recommend conducting testing with a more diverse user group, with a specific focus on accessibility and cognitive accessibility. Additionally, adopting a broader test setup that encompasses a wider range of symptoms and includes other eHealth services intended for broad usage would be beneficial.

Comparison With Prior Work

Usability issues were efficiently identified during testing with special user groups. In a study by Liu et al [36], which involved 350 participants, similar problems were discovered with service A as found in our study. The authors observed comparable challenges related to understanding questions and terminology, along with a need to enhance the visual layout and instructions for users. However, a notable disparity was observed in completion times: their participants completed the symptom checkers in an average of 4 minutes and 9 seconds, whereas users in our study required, on average, 3 times longer. In addition to uncovering issues that notably impact cognitive accessibility, our study identified similar usability problems as other assessments. Furthermore, as highlighted by Jormanainen et al [23], the same service was used over 1.5 million times for COVID-19 evaluation, suggesting its successful use by a vast number of users. Moreover, challenges with terminology have been recognized in other services [20].

This study has concentrated on cognitive accessibility with 3 distinct user groups. Comparable user groups have been used in other studies that center on eHealth services [1,29,30]. Upon comparing our findings with these studies, we observe that the necessity for clearer language and terminology, along with the clarity of the service, has previously been recognized through interviews and focus groups [1,29]. Our study provides more nuanced insights into how these issues manifest in practical usage.

Conclusions

In this study, we conducted a qualitative usability evaluation of 2 online symptom checkers, with a particular emphasis on the cognitive accessibility of the services. The evaluation targeted potentially vulnerable groups at risk of digital exclusion. Three distinct user groups participated in the tests: older adults, individuals with mild intellectual disabilities, and nonnative Finnish speakers. Our findings revealed that these groups encountered numerous difficulties with the tested services, particularly concerning their clarity and the language/terminology used. Furthermore, when testing with these groups, several key points must be considered: test sessions should be meticulously planned, instructions need to be clear, sessions should not be overly prolonged, and sufficient time must be allocated for each task.

In general, we found that testing with vulnerable groups was both useful and efficient. The rate of usability problems identified was notably high compared with the number of participants, and these issues were readily uncovered. These user groups encountered similar challenges related to information processing. It is imperative to provide them with better support through services that are clear, presenting less information and fewer options at once, and incorporating fewer long and complex words and selection lists. Additionally, following the principles of universal design, the proposed improvements are such that they will also benefit a more general user group. Therefore, we highly recommend testing with potentially vulnerable groups and, furthermore, expanding the user groups to include a representation of a broader variety of cognitive characteristics and challenges.

XSL•FO
RenderX

Acknowledgments

This research was conducted as part of the “Toward a Socially Inclusive Digital Society: Transforming Service Culture” project (DigiIN) and was funded by the Strategic Research Council at the Academy of Finland (grant numbers 352501and 352503) and NordForsk (project 100477). We thank the service providers DigiFinland Oy and Klinik Healthcare Solutions for allowing us to test the services, and Selkeästi meille initiative (a joint project between The Finnish Association on Intellectual and Developmental Disabilities [FAIDD] and Kehitysvammatuki 57 ry) and Väylä ry for helping us to contact the test users. Finally, we thank Felipe González Carceller for his excellent contribution to the research.

Data Availability

The data sets generated during or analyzed during this study are not publicly available due to the sensitive nature of the data but the numeric data are available from the corresponding author on reasonable request.

Conflicts of Interest

None declared.

Multimedia Appendix 1
The distribution of the symptoms and services.
[DOCX File , 17 KB - humanfactors_v11i1e45275_app1.docx]

Multimedia Appendix 2
Symptom vignettes used.
[DOCX File , 16 KB - humanfactors_v11i1e45275_app2.docx]

Multimedia Appendix 3
The background information of the participants. All the participants used digital services multiple times a day. MID: mildly intellectually disabled.
[DOCX File , 15 KB - humanfactors_v11i1e45275_app3.docx]

Multimedia Appendix 4
Usability problem categories.
[DOCX File , 15 KB - humanfactors_v11i1e45275_app4.docx]

Multimedia Appendix 5
Average SUS scores for both services. For older adults the two services got the same results, for the two other groups Service A got a slightly better score.
[DOCX File , 13 KB - humanfactors_v11i1e45275_app5.docx]

References

1. Kaihlanen A, Virtanen L, Buchert U, Safarov N, Valkonen P, Hietapakka L, et al. Towards digital health equity - a qualitative study of the challenges experienced by vulnerable groups in using digital health services in the COVID-19 era. BMC Health Serv Res 2022 Feb 12;22(1):188 [FREE Full text] [doi: 10.1186/s12913-022-07584-4] [Medline: 35151302]
2. European Commission. European accessibility act. European Commission. URL: https://ec.europa.eu/social/main.jsp?catId=1202 [accessed 2022-10-28]
3. Maramba I, Chatterjee A, Newman C. Methods of usability testing in the development of eHealth applications: a scoping review. Int J Med Inform 2019 Jun;126:95-104. [doi: 10.1016/j.ijmedinf.2019.03.018] [Medline: 31029270]
4. Archer N, Keshavjee K, Demers C, Lee R. Online self-management interventions for chronically ill patients: cognitive impairment and technology issues. Int J Med Inform 2014 Apr;83(4):264-272 [FREE Full text] [doi: 10.1016/j.ijmedinf.2014.01.005] [Medline: 24507762]
5. Story MF, Mueller JL, Mace RL. Designing for People of All Ages and Abilities. Raleigh, NC: The Center for Universal Design, NC State University; 1998.
6. International Organization for Standardization (ISO). Ergonomics of human-system interaction: Part 11: Usability: Definitions and concepts (ISO 9241-11:2018). International Organization for Standardization (ISO). 2018. URL: https://www.iso.org/standard/63500.html#:~:text=ISO%209241%2D11%3A2018%20provides,services%20(including%20technical%20and%20personal [accessed 2024-02-24]

XSL•FO
RenderX

7. International Organization for Standardization (ISO). ISO/IEC 25010:2019: Systems and software engineering -- Systems and software Quality Requirements and Evaluation (SQuaRE) -- System and software quality models. International Organization for Standardization (ISO). URL: https://iso25000.com/index.php/en/iso-25000-standards/iso-25010 [accessed 2024-02-24]
8. Bernard R, Sabariego C, Cieza A. Barriers and facilitation measures related to people with mental disorders when using the web: a systematic review. J Med Internet Res 2016 Jun 09;18(6):e157 [FREE Full text] [doi: 10.2196/jmir.5442] [Medline: 27282115]
9. Muñoz-Hernández S, Benac-Earle C, Herranz Nieva A, Gonzalez-McGuinness M. Sustainable methodology for operational and formal digital skills acquisition: a case study of e-health inclusion. Sustainability 2021 Aug 30;13(17):9698. [doi: 10.3390/su13179698]
10. Miesenberger K, Edler C, Heumader P, Petz A. In: Yesilada Y, Harper S, editors. Tools and Applications for Cognitive Accessibility. London, UK: Springer; 2019:523-546.
11. Chambers D, Cantrell AJ, Johnson M, Preston L, Baxter SK, Booth A, et al. Digital and online symptom checkers and health assessment/triage services for urgent health problems: systematic review. BMJ Open 2019 Aug 01;9(8):e027743 [FREE Full text] [doi: 10.1136/bmjopen-2018-027743] [Medline: 31375610]
12. Meyer AND, Giardina TD, Spitzmueller C, Shahid U, Scott TMT, Singh H. Patient perspectives on the usefulness of an artificial intelligence-assisted symptom checker: cross-sectional survey study. J Med Internet Res 2020 Jan 30;22(1):e14679 [FREE Full text] [doi: 10.2196/14679] [Medline: 32012052]
13. Marco-Ruiz L, Bønes E, de la Asunción E, Gabarron E, Aviles-Solis JC, Lee E, et al. Combining multivariate statistics and the think-aloud protocol to assess Human-Computer Interaction barriers in symptom checkers. J Biomed Inform 2017 Oct;74:104-122 [FREE Full text] [doi: 10.1016/j.jbi.2017.09.002] [Medline: 28893671]
14. Alajarmeh N. Evaluating the accessibility of public health websites: an exploratory cross-country study. Univers Access Inf Soc 2022 Jan 27;21(3):771-789 [FREE Full text] [doi: 10.1007/s10209-020-00788-7] [Medline: 33526996]
15. Helsper EJ. A corresponding fields model for the links between social and digital exclusion. Commun Theor 2012 Oct 15;22(4):403-426. [doi: 10.1111/j.1468-2885.2012.01416.x]
16. Mulvale G, Moll S, Miatello A, Robert G, Larkin M, Palmer VJ, et al. Codesigning health and other public services with vulnerable and disadvantaged populations: insights from an international collaboration. Health Expect 2019 Jun;22(3):284-297 [FREE Full text] [doi: 10.1111/hex.12864] [Medline: 30604580]
17. Arsenijevic J, Tummers L, Bosma N. Adherence to electronic health tools among vulnerable groups: systematic literature review and meta-analysis. J Med Internet Res 2020 Feb 06;22(2):e11613 [FREE Full text] [doi: 10.2196/11613] [Medline: 32027311]
18. Fang ML, Siden E, Korol A, Demestihas M, Sixsmith J, Sixsmith A. A scoping review exploration of the intended and unintended consequences of eHealth on older people: a health equity impact assessment. Human Technology 2018 Nov 30:297-323. [doi: 10.17011/ht/urn.201811224835]
19. Kampmeijer R, Pavlova M, Tambor M, Golinowska S, Groot W. The use of e-health and m-health tools in health promotion and primary prevention among older adults: a systematic literature review. BMC Health Serv Res 2016 Sep 05;16 Suppl 5(Suppl 5):290 [FREE Full text] [doi: 10.1186/s12913-016-1522-3] [Medline: 27608677]
20. Sayago S, Blat J. Telling the story of older people e-mailing: an ethnographical study. International Journal of Human-Computer Studies 2010 Jan;68(1-2):105-120 [FREE Full text] [doi: 10.1016/j.ijhcs.2009.10.004]
21. Nunes F, Silva PA, Abrantes F. Human-Computer Interaction and the older adult: an example using user research and personas. In: Proceedings of the 3rd International Conference on Pervasive Technologies Related to Assistive Environments, PETRA 2010, Samos, Greece, June 23-25, 2010. 2010 Presented at: The 3rd International Conference on Pervasive Technologies Related to Assistive Environments (PETRA 2010); June 23-25, 2010; Samos, Greece p. 1-8 URL: https://www.researchgate.net/publication/221410323_Human-Computer_Interaction_and_the_older_adult_An_example_using_user_research_and_personas [doi: 10.1145/1839294.1839353]
22. Satre DD, Knight BG, David S. Cognitive-behavioral interventions with older adults: integrating clinical and gerontological research. Professional Psychology: Research and Practice 2006 Oct;37(5):489-498. [doi: 10.1037/0735-7028.37.5.489]
23. Jormanainen V, Soininen L. Use and users of the web-based Omaolo COVID-19 symptom self-assesment tool in Finland since March 16, 2020. Stud Health Technol Inform 2021 May 27;281:739-743. [doi: 10.3233/SHTI210270] [Medline: 34042674]
24. Lussier-Desrochers D, Normand CL, Romero-Torres A, Lachapelle Y, Godin-Tremblay V, Dupont M, et al. Bridging the digital divide for people with intellectual disability. CP 2017 May 31;11(1):Article 1. [doi: 10.5817/cp2017-1-1]
25. Alfredsson Ågren K, Kjellberg A, Hemmingsson H. Digital participation? Internet use among adolescents with and without intellectual disabilities: a comparative study. New Media & Society 2019 Nov 18;22(12):2128-2145. [doi: 10.1177/1461444819888398]
26. Marston HR, Genoe R, Freeman S, Kulczycki C, Musselwhite C. Older adults' perceptions of ICT: main findings from the Technology In Later Life (TILL) study. Healthcare (Basel) 2019 Jul 04;7(3):7030086 [FREE Full text] [doi: 10.3390/healthcare7030086] [Medline: 31277387]

XSL•FO
RenderX

27. Setchell J, Barlott T, Torres M. A socio-emotional analysis of technology use by people with intellectual disabilities. J Intellect Disabil Res 2021 Feb 22;65(2):149-161. [doi: 10.1111/jir.12796] [Medline: 33225540]
28. Fernández Da Silva Á, Buceta BB, Mahou-Lago XM. eHealth policy in Spain: a comparative study between general population and groups at risk of social exclusion in Spain. Digit Health 2022;8:20552076221120724 [FREE Full text] [doi: 10.1177/20552076221120724] [Medline: 35990110]
29. Oudshoorn CEM, Frielink N, Nijs SLP, Embregts PJCM. eHealth in the support of people with mild intellectual disability in daily life: a systematic review. J Appl Res Intellect Disabil 2020 Nov 23;33(6):1166-1187 [FREE Full text] [doi: 10.1111/jar.12758] [Medline: 32578361]
30. Ware P, Bartlett SJ, Paré G, Symeonidis I, Tannenbaum C, Bartlett G, et al. Using eHealth technologies: interests, preferences, and concerns of older adults. Interact J Med Res 2017 Mar 23;6(1):e3 [FREE Full text] [doi: 10.2196/ijmr.4447] [Medline: 28336506]
31. Kadir SA, Jamaludin M. Universal design as a significant component for sustainable life and social development. Procedia - Social and Behavioral Sciences 2013 Sep;85:179-190 [FREE Full text] [doi: 10.1016/j.sbspro.2013.08.349]
32. Ruzic L, Sanfod J. Universal design mobile interface guidelines (UDMIG) for an aging population. In: Marston HR, Freeman S, Musselwhite C, editors. Mobile e-Health. Cham, Switzerland: Springer International Publishing; 2017:17-37.
33. Raspa M, Fitzgerald T, Furberg R, Wylie A, Moultrie R, DeRamus M, et al. Mobile technology use and skills among individuals with fragile X syndrome: implications for healthcare decision making. J Intellect Disabil Res 2018 Oct;62(10):821-832 [FREE Full text] [doi: 10.1111/jir.12537] [Medline: 30105880]
34. Watfern C, Heck C, Rule C, Baldwin P, Boydell KM. Feasibility and acceptability of a mental health website for adults with an intellectual disability: qualitative evaluation. JMIR Ment Health 2019 Mar 28;6(3):e12958 [FREE Full text] [doi: 10.2196/12958] [Medline: 30920378]
35. Persson H, Åhman H, Yngling AA, Gulliksen J. Universal design, inclusive design, accessible design, design for all: different concepts—one goal? On the concept of accessibility—historical, methodological and philosophical aspects. Univ Access Inf Soc 2014 May 7;14(4):505-526. [doi: 10.1007/s10209-014-0358-z]
36. Liu VD, Sellgren L, Kaila M, Koskela T. Sähköisten oirearvioiden käytettävyys. FinJeHeW 2021 Jun 18;13(2):97020. [doi: 10.23996/fjhw.97020]
37. Kim H. A study on usability improvement of mobile health care services. Int J Adv Smart Converg 2017;6(2):72-81. [doi: 10.7236/IJASC.2017.6.2.72]
38. Knitza J, Mohn J, Bergmann C, Kampylafka E, Hagen M, Bohr D, et al. Accuracy, patient-perceived usability, and acceptance of two symptom checkers (Ada and Rheport) in rheumatology: interim results from a randomized controlled crossover trial. Arthritis Res Ther 2021 Apr 13;23(1):112 [FREE Full text] [doi: 10.1186/s13075-021-02498-8] [Medline: 33849654]
39. Broekhuis M, van Velsen L, Peute L, Halim M, Hermens H. Conceptualizing usability for the eHealth context: content analysis of usability problems of ehealth applications. JMIR Form Res 2021 Jul 27;5(7):e18198 [FREE Full text] [doi: 10.2196/18198] [Medline: 34313594]
40. Omaolo. DigiFInland Oy. URL: https://www.omaolo.fi [accessed 2022-07-10]
41. Klinik Access. Klinik Healthcare Solutions Oy. URL: http://www.klinik.fi [accessed 2022-07-10]
42. Etikan I. Comparison of convenience sampling and purposive sampling. AJTAS 2016;5(1):1. [doi: 10.11648/j.ajtas.20160501.11]
43. Caine K. Local standards for sample size at CHI. In: CHI '16: Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems. New York, NY: Association for Computing Machinery; 2016 May Presented at: CHI '16: The 2016 CHI Conference on Human Factors in Computing Systems; May 7-12, 2016; San Jose, CA p. 981-992. [doi: 10.1145/2858036.2858498]
44. Semigran H, Linder J, Gidengil C, Mehrotra A. Evaluation of symptom checkers for self diagnosis and triage: audit study. BMJ 2015 Jul 08;351:h3480. [doi: 10.1136/bmj.h3480] [Medline: 26157077]
45. North F, Ward WJ, Varkey P, Tulledge-Scheitel SM. Should you search the internet for information about your acute symptom? Telemed J E Health 2012 Apr;18(3):213-218. [doi: 10.1089/tmj.2011.0127] [Medline: 22364307]
46. Eronen J, Paakkari L, Portegijs E, Saajanaho M, Rantanen T. Assessment of health literacy among older Finns. Aging Clin Exp Res 2019 Apr 21;31(4):549-556 [FREE Full text] [doi: 10.1007/s40520-018-1104-9] [Medline: 30578457]
47. Brooke J. SUS: a 'quick and dirty' usability scale. In: Usability Evaluation In Industry (1st Edition). London, UK: CRC Press; 1996:1-6.
48. Bangor A, Kortum PT, Miller JT. An Empirical Evaluation of the System Usability Scale. International Journal of Human-Computer Interaction 2008 Jul 30;24(6):574-594. [doi: 10.1080/10447310802205776]

Abbreviations

SUS: System Usability Scale
UI: user interface

XSL•FO
RenderX

Edited by Y Gong; submitted 22.12.22; peer-reviewed by C Preiser, E Bønes; comments to author 16.02.23; revised version received 08.03.23; accepted 03.02.24; published 08.03.24.

Please cite as:
Savolainen K, Kujala S
Testing Two Online Symptom Checkers With Vulnerable Groups: Usability Study to Improve Cognitive Accessibility of eHealth Services
JMIR Hum Factors 2024;11:e45275
URL: https://humanfactors.jmir.org/2024/1/e45275
doi:10.2196/45275
PMID:38457214

©Kaisa Savolainen, Sari Kujala. Originally published in JMIR Human Factors (https://humanfactors.jmir.org), 08.03.2024. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on https://humanfactors.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Head Protection Device for Individuals at Risk for Head Injury due to Ground-Level Falls: Single Trauma Center User Experience Investigation

Susan Haag[1*], PhD; John Kepros[1*], MD

Scottsdale Osborn Medical Center, Scottsdale, AZ, United States

[*]all authors contributed equally

Corresponding Author:
Susan Haag, PhD
Scottsdale Osborn Medical Center
7400 E Osborn Rd
Scottsdale, AZ, 85251
United States
Phone: 1 480 323 4018
Email: susan.haag@ymail.com

Abstract

Background: Falls represent a large percentage of hospitalized patients with trauma as they may result in head injuries. Brain injury from ground-level falls (GLFs) in patients is common and has substantial mortality. As fall prevention initiatives have been inconclusive, we changed our strategy to injury prevention. We identified a head protection device (HPD) with impact-resistant technology, which meets head impact criteria sustained in a GLF. HPDs such as helmets are ubiquitous in preventing head injuries in sports and industrial activities; yet, they have not been studied for daily activities.

Objective: We investigated the usability of a novel HPD on patients with head injury in acute care and home contexts to predict future compliance.

Methods: A total of 26 individuals who sustained head injuries, wore an HPD in the hospital, while ambulatory and were evaluated at baseline and 2 months post discharge. Clinical and demographic data were collected; a usability survey captured HPD domains. This user experience design revealed patient perceptions, satisfaction, and compliance. Nonparametric tests were used for intragroup comparisons (Wilcoxon signed rank test). Differences between categorical variables including sex, race, and age (age group 1: 55-77 years; age group 2: 78+ years) and compliance were tested using the chi-square test.

Results: Of the 26 patients enrolled, 12 (46%) were female, 18 (69%) were on anticoagulants, and 25 (96%) were admitted with a head injury due to a GLF. The median age was 77 (IQR 55-92) years. After 2 months, 22 (85%) wore the device with 0 falls and no GLF hospital readmissions. Usability assessment with 26 patients revealed positive scores for the HPD post discharge regarding satisfaction (mean 4.8, SD 0.89), usability (mean 4.23, SD 0.86), effectiveness (mean 4.69, SD 0.54), and relevance (mean 4.12, SD 1.10). Nonparametric tests showed positive results with no significant differences between 2 observations. One issue emerged in the domain of aesthetics; post discharge, 8 (30%) patients had a concern about device weight. Analysis showed differences in patient compliance regarding age (χ_1^2=4.27; *P*=.04) but not sex (χ_1^2=1.58; *P*=.23) or race (χ_1^2=0.75; *P*=.60). Age group 1 was more likely to wear the device for normal daily activities. Patients most often wore the device ambulating, and protection was identified as the primary benefit.

Conclusions: The HPD intervention is likely to have reasonably high compliance in a population at risk for GLFs as it was considered usable, protective, and relevant. The feasibility and wearability of the device in patients who are at risk for GLFs will inform future directions, which includes a multicenter study to evaluate device compliance and effectiveness. Our work will guide other institutions in pursuing technologies and interventions that are effective in mitigating injury in the event of a fall in this high-risk population.

(JMIR Hum Factors 2024;11:e54854) doi:10.2196/54854

XSL•FO
RenderX

KEYWORDS

health care interventions and technologies; user experience research; usability; brain injury; ground-level fall (GLF); head protection device (HPD); fall risk; patient compliance

Introduction

Frailty in aging is represented by a decline in functioning, with a risk of poor outcomes, including falls, which have implications for clinical practice and public health [1]. Falls are the primary cause of injury-related death in aging adults as 33% of adults 65 years and older fall each year [2,3].

Falls also represent a large percentage of hospitalized older patients as they may result in multiple injuries, including head trauma [4-6]. A head injury can be a common cause of disability and mortality and may be as mild as a bump, bruise (contusion), or cut and can be moderate to severe due to a concussion. Head injury may lead to premature nursing home admissions and increased hospital length of stay (LOS) with undesirable results for patients and hospitals [7,8]. Due to the aging population worldwide, the incidence of falls will continue to rise [9,10].

Studies have shown a clear pattern of increased health care costs associated with falls and frail individuals and various fall prevention initiatives have been promoted. Of the fall prevention interventions studied, some results have been favorable, such as those with well-developed educational programs [11]. However, others have been inconclusive [12-14], prompting our center to include head injury prevention, and therefore, we investigated a head protection device (HPD), similar to a helmet. In many fields, such as construction and sports, helmets have shown efficacy in preventing head injury risks, especially moderate to severe head injury [15-17]. The human head is vulnerable to even moderate impact as it can cause injury or death. A greater emphasis has been placed on job safety in industries like construction particularly to protect the head from injury, and hard hats and helmets have been required [18]. However, historically, helmets have not been used for normal daily living.

Health care systems are increasingly looking for contexts that provide accessible and efficient care and for medical devices and interventions to improve the patient experience and health outcomes [19,20]. Human factors, a scientific discipline, is important in clinical practice as it reveals how humans interact with interventions, such as devices, regarding expectations and limitations. User experience (UX) focuses on having a deep understanding of users and what they need and value [21-23]. UX research has been used to ascertain user domains such as adherence, usability, and perceived impact and has assisted with intervention development and refinement [24]. Adopting a UX research design will help ensure that new devices are easy to use and meet the needs of most patients.

Clinical practices should target effective strategies that improve individuals' quality of life and independence including screenings and interventions to manage injuries associated with falls [25,26]. Screenings that measure activities of daily living (ADLs) are essential, as the ability to perform daily tasks safely without exhaustion is a critical component of healthy aging, thus allowing older individuals to maintain their independence and quality of life [27]. Measurement of daily activities is important as these may be predictors of early admission to assisted care facilities or the need for alternative living arrangements [28,29].

Recent literature advocates change toward tailored interventions that preserve an individual's independence by promoting furthering advancements in evidence-based treatment options and identifying cost-effective strategies [2,3]. Due to an increasing incidence of head injuries after ground-level falls (GLFs) in our trauma center, we designed a study that examined the effects of a low-cost HPD that has the potential to prevent head injury due to a fall.

The purpose of this UX research was to assess compliance by investigating the usability of an HPD from a patient's perspective in both acute care (hospital) and home contexts. We hypothesized that consented patients would follow the research protocol as recommended and wear the device in the hospital and at the 2 months post discharge. The primary limitation in an aging population is compliance, which we approached first. This in-hospital and home-based UX investigation concerning a low-cost treatment option may serve clinicians to better manage frailty and mitigate injury due to falls in their clinical practice.

Methods

Study Design

We considered the UX of frail individuals at this developmental, exploratory stage of a device to examine patient adherence and use. The UX assessment instrument adopted UX domains with a 5-point scale showing a more positive rating (rating of 5) and a lower rating (rating of 1). UX domains included device credibility, satisfaction, usability, adherence, effectiveness, relevance, and aesthetics. The primary outcome variable is patient compliance regarding wearing the device for 2 months. Additional data collected included the frequency of wearing the device during normal daily activities. Consistent with the literature, ADLs (such as ambulating and preparing meals) are critical for independence in aging populations [29].

Recruitment

Participants were recruited from among patients who were treated at our level 1 trauma center and subsequently admitted to the hospital for observation due to head injury. Protocol inclusion criteria included the following: patients admitted to the hospital with a fall sustaining a head injury, patients with fall risk (eg, patients who fell within the prior year or other physical conditions aligned with fall risk), and patients who were ambulatory and 55 years or older. Head injuries included in the study were patients with a concussion, contusion, lacerations, or loss of consciousness. The individuals recruited did not experience trauma that required surgical intervention. After signing the consent in the hospital, individuals were given an HPD at no cost to wear while ambulatory. After consenting

XSL•FO
RenderX

and wearing the HPD for in-hospital observation (and just before discharge), the hospital team asked whether the patients would wear the HPD at home. If the patient agreed, we indicated that the research team would follow up post discharge for additional observations using the UX survey.

Ethical Considerations

In total, 26 patients, who experienced a fall and sustained a head injury, wore an HPD in hospital, while ambulatory and were evaluated at baseline (before discharge) and at 2 months post discharge. The study protocol was approved by the institutional review board for research ethics and subsequently approved (IRB 1804935). Informed consent was obtained from the 26 patients who met the inclusion criteria and were willing to participate. Confidentiality of information was maintained. The data are anonymized and patients are deidentified. Each patient was assigned a discrete number in the study and data are secured by the research scientist. There was no compensation for patient participation in the study.

HPD

The HPD includes an impact-resistant technological insert for additional head protection. It helps protect against bumps, scrapes, bruises, and other head injuries. The HPD is designed with ventilation to provide airflow for breathability without compromised protection. The HPD size can be adjusted with a hook and loop strap to give a quick, secure fit. Figure 1 displays the HPD, which looks like a typical baseball cap.

Figure 1. Head protection device.

Usability Survey

A multidisciplinary health care team comprised of physicians, a research scientist, and physical therapists collaborated on the study design, developing a usability survey for patients who are at risk of fall, which led to a tangible and targeted intervention strategy. UX (usability) domain definitions were identified in the literature. Existing domain definitions were examined such as credibility, usability, and satisfaction [24], and additional domains were defined such as effectiveness, relevance, and aesthetics. The domains were refined, used on the usability survey instrument, and functioned as outcome measures. Textbox 1 shows the domains and UX definitions. UX domain data were collected on the instrument using a 5-point scale (5=strongly agree, 4=agree, 3=neutral, 2=disagree, and 1=strongly disagree). Patients were asked if they would recommend the HPD. The survey was intended to evaluate the HPD's usability and was administered after patients concluded their interaction with the HDP in the hospital. Those who agreed to wear the HPD at home were provided a device and were reevaluated post discharge.

Textbox 1. Domain and user experience definitions.

- Credibility: whether the user perceives the device to be trustworthy (eg, accuracy and quality of information presented in the patient consent)
- Satisfaction: the user's overall experience and interaction with the device
- Usability: the user's perceived ease of use of the device based on technical factors
- Adherence: whether the patient followed the device research protocol and continued to use the device as recommended (compliance) completing outcome measures
- Effectiveness: the extent the user perceives the overall value of the device, including safety and whether they would recommend it to another fall risk individual
- Relevance: the extent to which the device is appropriate for their situation and whether they perceive it meets their needs (provides protection to their head and helps them maintain a sense of independence)
- Aesthetics: factors such as color, pattern, size, shape, and weight

XSL•FO
RenderX

Data Collection

Quantitative data included demographics (age, sex, and race) and clinical data such as hospital LOS, number of GLFs, readmission to the hospital due to a GLF, and Glasgow Coma Scale. Data were also captured on the usability survey including domains such as device satisfaction, effectiveness, relevance, and aesthetics. Qualitative data were also collected on the usability survey, and patient comments were recorded regarding HPD benefits and opportunities for improvement.

Statistical Analysis

This UX research methodology included multiple patient observations and differences between observations were examined. Nonparametric tests, used to analyze ordinal and categorical data, were used for intragroup comparisons (Wilcoxon signed rank test). We used descriptive statistics, such that patterns might emerge from the data. Frequencies and percentages are reported for categorical variables. Medians and means with SDs are reported for continuous variables as appropriate. All computations included 26 patients. Group comparisons were made using chi-square tests or Fisher exact tests, where numbers were small and were reported as numbers (%). All variables were assessed for normality. Analyses of categorical variables (age) and patient adherence were tested using the chi-square statistic. Statistical tests are 2-tailed, with a significance level of an α of .05. All statistical analyses were performed using SPSS Statistics for Windows (version 28.0; IBM Corp).

Open-ended patient comments (qualitative data) were analyzed using a 3-step process: data reduction, data display, and conclusion drawing and verification. Data reduction helped sort and compile data excerpts (to organize the data) and assist in developing assertions regarding patient perceptions surrounding wearability (eg, comfort and weight) and modifications of HPD, if necessary. Excerpts were annotated with topics such as the benefits of HPD: positive feedback (aspects recorded as positive by the patient participants regarding HPD experience and interaction) and negative feedback (points considered negative by the patients pertaining to interaction with the device). As a next step, we analyzed the themes that emerged and categorized them based on whether they were related to the usability of the HPD or the health support the device offered.

Results

Study Population

Among the 26 participants, 12 (46%) were female and 5 (19%) were non-White, with a median age of 77 (IQR 55-92) years. The average hospital LOS was 3.8 (SD 3.65) days. The majority (n=25, 96%) of patients who experienced head trauma were admitted to the hospital with a head injury due to a GLF (n=1, 4% were other types of falls); 22 (85%) had prior falls in the last 12 months and 16 (62%) had a hospital visit due to a head injury related to a fall within the year; 18 (69%) were on anticoagulants. The mean Glasgow Coma Score was 14.2 (SD 0.44). The age category was divided into 2 groups for analysis: age group 1 comprised of those who were 55 to 77 years and age group 2 comprised of patients 78 years and older.

Usability Survey Domain Results

In the hospital, all 26 consented patients wore the device with 0 falls recorded. After 2 months, 22 (85%) were wearing the HPD, had 0 falls, and had no hospital readmissions due to GLFs. At 6 months, 16 (62%) patients were compliant with wearing the device, with 0 falls and no hospital readmissions due to a GLF. The results showed positive scores, with no significant differences between ratings in hospital and post discharge regarding device credibility (0.42), satisfaction (0.60), usability (0.80), adherence (0.06), effectiveness (0.53), and relevance (0.09). A difference emerged for the domain of aesthetics. After the discharge, 8 (30%) patients had concerns regarding the device's weight, saying it was slightly heavier than a typical cap. Overall, users had a positive experience with the HPD and scores revealed that patients felt it was effective and relevant. Thus, post discharge, users would recommend the HPD to others at risk for falls (mean 4.52, SD 0.51). Users were compliant by wearing the device in hospital and at 2 months post discharge, supporting the research hypothesis. Table 1 displays the UX domain means (SDs) for 2 observations.

Differences between categorical variables (age group 1: 55-77 years, group 2: 78 years and older, sex, and race) and protocol adherence were analyzed. Chi-square analysis showed differences in compliance regarding age (χ_1^2=4.27; *P*=.04) but not sex (χ_1^2=1.58; *P*=.23) or race (χ_1^2=0.75; *P*=.60). Age group 1 was more likely to wear the device for normal daily activities.

Table 1. In-hospital and postdischarge intragroup domain differences.

User experience domains	Hospital, mean (SD)	Postdischarge, mean (SD)	*P* value
Credibility	3.91 (0.80)	4.01 (0.84)	.42
Satisfaction	4.15 (0.88)	4.80 (0.89)	.60
Usability	4.27 (0.66)	4.23 (0.86)	.80
Adherence	4.50 (0.86)	4.30 (1.06)	.06
Effectiveness (value)	4.62 (0.49)	4.69 (0.54)	.53
Relevance	4.42 (0.75)	4.12 (1.10)	.09
Aesthetics	3.38 (1.30)	2.96 (1.83)	.003

XSL•FO
RenderX

Patient Device Use in Daily Activities

The usability survey data captured patient device use during typical ADLs at 2 weeks and at 2 months post discharge. Users were provided a list of daily activities and were asked to rate the frequency of wearing the device. Consistent with the literature, ADLs, such as ambulating and preparing meals, are critical for independence in an aging population [29]. The highest score on the usability instrument was a "5" which indicated that the patient would wear the HPD "most often." In-home contexts, patients indicated they most often wore the device ambulating and when driving (to meals and doctor appointments) and less often for personal hygiene. Table 2 shows within-group differences in device use in daily activities.

Table 2. Within-group differences in device use in daily activities.

Daily activities	Two weeks, mean (SD)	Two months, mean (SD)	*P* value
Ambulating	4.31 (0.92)	4.15 (1.12)	.47
Driving (or being driven)	4.04 (0.77)	4.12 (0.76)	.16
Grocery shopping or shopping	3.69 (1.28)	3.58 (1.23)	.54
Relaxing (TV)	4.00 (1.06)	3.31(1.10)	.20
Housekeeping	3.35 (1.09)	3.27 (1.00)	.67
Preparing meals	2.77 (0.99)	2.50 (1.06)	.07
Personal hygiene	2.42 (0.94)	2.27 (0.96)	.49

Positive Patient Feedback

Open-ended questions on the usability instrument elicited patient qualitative comments regarding HPD benefits and opportunities for improvement. As a result, 2 dominant themes emerged, namely HPD usability and HPD as health support (protection). Usability was associated with the use of the device and functionality in terms of wearability. Health support included themes that were aligned with head protection for a patient.

Usability and relevance from the patients' perspective translated into wearability, and the majority of patients wore the device after 2 months post discharge. Participants felt that the HPD was comfortable and easy to wear. However, 8 (30%) patients mentioned that the HPD was not as light as a typical cap due to the protective "technology insert" and suggested the HPD could be lighter in weight. One male participant stated,

> *The cap is heavier than a usual baseball cap and it took me longer to get used to it. I would like it a bit lighter in weight if possible and more air vents to let in air.*

Health support from the participant's perspective sufficed as the primary benefit, as 18 (69%) commented that the device protected their head in the event of a fall. Patients called the device a "cap" as it resembles a baseball cap. One patient stated, "Protection for my head is important. I will wear it going out to eat and to doctor appointments." Another female participant indicated, "I wear it eight hours a day to protect my head." Two patients (male and female) indicated post discharge, they hit their heads on cabinets, as 1 commented:

> *I already bent over and hit my head on a cabinet; it protected me from another head injury. Since wearing the cap, I have not had a fall, only a bump and I had on my cap.*

A 74-year-old female participant stated, "I fell last year and I will wear this walking whenever possible. It protects my head." A male participant noted, "The device is protective and comfortable; I forgot I had it on." From patient comments, the HPD is cognate with head protection.

Discussion

Principal Findings

Using a UX design, we investigated the usability of a novel HPD on patients with head injury in acute care and home contexts to predict future compliance. All 26 patients provided positive scores for the HPD post discharge regarding satisfaction, usability, effectiveness, and relevance. Nonparametric tests showed positive results, with no significant differences between 2 observations at 2 months. Chi-square analysis showed a significant difference in HPD compliance regarding age but not sex or race as age group 1 was more likely to wear the device for normal daily activities. Patients most often wore the device ambulating and head protection was identified as the primary benefit. Thus, patients were most likely to recommend the HPD to others at risk of GLFs.

Due to the consistently high rate of head injuries after GLFs in our center, the targeted team strategy for an HPD and UX research design was developed. We realized that patient compliance in the geriatric population has been a limiting factor and approached that aspect first. Patients adhered to the research protocol by wearing the device in the hospital and post discharge, in the home, supporting the research hypothesis. At 2 months, 22 (85%) patients wore the device with 0 falls recorded and no readmissions due to falls.

Our multidisciplinary team, a diverse group of medical professionals, consisting of physicians, research scientists, and physical therapists, studied a device to be worn during daily activities in home environments. Recent literature has advocated for home care strategies [30] and interventions to be used in home contexts where falls most often occur [31]. Managing falls in this high-risk population is complex, requiring a systemic and collaborative approach directed by a multidisciplinary team focused on improving patient outcomes [3].

XSL•FO
RenderX

Limitations

Accuracy is critical regarding the collection of patient data, and the in-hospital data collection was conducted under medical supervision. However, the limitations of the UX research included the nature of self-reporting by participants post discharge at 2 and 6 months. One measure to counter this bias was to include a family member during the evaluation to corroborate the patient's self-reported data and responses. Another issue and limitation, we noted, was the difficulty of trying to reconnect or contact this population at follow-up due to cognitive decline, the extent and severity of head trauma, and other injuries associated with a GLF.

Conclusions

The results show our proposed HPD intervention will have a high compliance rate in those at risk for GLFs as it was considered usable, protective, and relevant. Managing individuals with fall risk may include future investigations of specific interventions and low-cost devices that preserve a patient's independence and physical function, and research that contributes to further advancements in evidence-based treatment options. The feasibility and wearability of the device in patients with GLF with head injuries will inform future directions, which includes a multicenter study to evaluate compliance and device effectiveness. Our work will guide other health care institutions in pursuing cost-effective treatments and technological interventions that are usable and effective in improving outcomes for this fall risk population.

Conflicts of Interest

None declared.

References

1. Dent E, Martin FC, Bergman H, Woo J, Romero-Ortuno R, Walston JD. Management of frailty: opportunities, challenges, and future directions. Lancet 2019;394(10206):1376-1386. [doi: 10.1016/S0140-6736(19)31785-4] [Medline: 31609229]
2. Hsieh KL, Fanning JT, Rogers WA, Wood TA, Sosnoff JJ. A fall risk mHealth app for older adults: development and usability study. JMIR Aging 2018 Jan 12;1(2):1-12 [FREE Full text] [doi: 10.2196/11569] [Medline: 31518234]
3. Jacobson CL, Foster LC, Arul H, Rees A, Stafford RS. A digital health fall prevention program for older adults: feasibility study. JMIR Form Res 2021 Dec 23;5(12):1-11 [FREE Full text] [doi: 10.2196/30558] [Medline: 34837492]
4. Dunnell L, Barnard AN, Chu K, Barling A, Birns J, Walker G. In situ simulation training for frailty. Geriatrics (Basel) 2023 Feb 17;8(1):1-9 [FREE Full text] [doi: 10.3390/geriatrics8010026] [Medline: 36826368]
5. Khurrum M, Chehab M, Ditillo M, Richards J, Douglas M, Bible L, et al. Trends in geriatric ground-level falls: report from the National Trauma Data Bank. J Surg Res 2021 May 24;266:261-268. [doi: 10.1016/j.jss.2021.02.047] [Medline: 34034061]
6. Tricco AC, Thomas SM, Veroniki AA, Hamid JS, Cogo E, Strifler L, et al. Comparisons of interventions for preventing falls in older adults: a systematic review and meta-analysis. JAMA 2017 Nov 07;318(17):1687-1699 [FREE Full text] [doi: 10.1001/jama.2017.15006] [Medline: 29114830]
7. Jones DM, Connelly DW, Burgess JR. Ground level falls in patients over 60, do race and socioeconomic status affect outcome? Am Surg 2022 Apr 23;88(9):2127-2131. [doi: 10.1177/00031348221091961] [Medline: 35465738]
8. Cameron ID, Dyer SM, Panagoda CE, Murray GR, Hill KD, Cumming RG, et al. Interventions for preventing falls in older people in care facilities and hospitals. Cochrane Database Syst Rev 2018 Sep 07;9(9):1-255 [FREE Full text] [doi: 10.1002/14651858.CD005465.pub4] [Medline: 30191554]
9. Grabel ZJ, Lunati MP, Segal DN, Kukowski NR, Yoon ST, Jain A. Thoracolumbar spinal fractures associated with ground level falls in the elderly: an analysis of 254,486 emergency department visits. J Clin Orthop Trauma 2020 Apr 10;11(5):916-920 [FREE Full text] [doi: 10.1016/j.jcot.2020.04.009] [Medline: 32879581]
10. Yee G, Jain A. Geriatric head injury. In: StatPearls. Treasure Island (FL): StatPearls Publishing; Jan 07, 2021:1-10.
11. Heng H, Jazayeri D, Shaw L, Kiegaldie D, Hill AM, Morris ME. Hospital falls prevention with patient education: a scoping review. BMC Geriatr 2020 Apr 15;20(1):140-147 [FREE Full text] [doi: 10.1186/s12877-020-01515-w] [Medline: 32293298]
12. Lamb SE, Bruce J, Hossain A, Ji C, Longo R, Lall R, et al. Screening and intervention to prevent falls and fractures in older people. N Engl J Med 2020 Nov 05;383(19):1848-1859 [FREE Full text] [doi: 10.1056/NEJMoa2001500] [Medline: 33211928]
13. Rezaei-Shahsavarloo Z, Atashzadeh-Shoorideh F, Gobbens RJJ, Ebadi A, Harouni GG. The impact of interventions on management of frailty in hospitalized frail older adults: a systematic review and meta-analysis. BMC Geriatr 2020 Dec 03;20(1):526-529 [FREE Full text] [doi: 10.1186/s12877-020-01935-8] [Medline: 33272208]
14. Williams CT, Whyman J, Loewenthal J, Chahal K. Managing geriatric patients with falls and fractures. Orthop Clin North Am 2023 May 31;54(3S):e1-e12. [doi: 10.1016/j.ocl.2023.04.001] [Medline: 37349065]
15. McIntosh AS, Andersen TE, Bahr R, Greenwald R, Kleiven S, Turner M, et al. Sports helmets now and in the future. Br J Sports Med 2011 Dec 04;45(16):1258-1265. [doi: 10.1136/bjsports-2011-090509] [Medline: 22117017]
16. Olivier J, Creighton P. Bicycle injuries and helmet use: a systematic review and meta-analysis. Int J Epidemiol 2017 Feb 01;46(1):278-292 [FREE Full text] [doi: 10.1093/ije/dyw153] [Medline: 27450862]

XSL•FO
RenderX

17. Long J, Yang J, Lei Z, Liang D. Simulation-based assessment for construction helmets. Comput Methods Biomech Biomed Engin 2015 Mar 18;18(1):24-37. [doi: 10.1080/10255842.2013.774382] [Medline: 23495784]
18. Wu JZ, Pan CS, Wimer BM, Rosen CL. Finite element simulations of the head-brain responses to the top impacts of a construction helmet: effects of the neck and body mass. Proc Inst Mech Eng H 2017 Dec 21;231(1):58-68 [FREE Full text] [doi: 10.1177/0954411916678017] [Medline: 28097935]
19. Bitkina OV, Jeong H, Lee BC, Park J, Park J, Kim HK. Perceived trust in artificial intelligence technologies: a preliminary study. Hum Factors Ergon Manuf Serv Ind 2020;30(4):282-290. [doi: 10.1002/hfm.20839]
20. Williams PAH, Lovelock B, Cabarrus JA. A sense of coherence approach to improving patient experience using information infrastructure modeling: design science research. JMIR Form Res 2022 Apr 12;6(4):e35418 [FREE Full text] [doi: 10.2196/35418] [Medline: 35307641]
21. Sauer J, Sonderegger A, Schmutz S. Usability, user experience and accessibility: towards an integrative model. Ergonomics 2020 Jun 12;63(10):1207-1220. [doi: 10.1080/00140139.2020.1774080] [Medline: 32450782]
22. Allen J, Hutchinson AM, Brown R, Livingston PM. User experience and care integration in transitional care for older people from hospital to home: a meta-synthesis. Qual Health Res 2017 Jul 28;27(1):24-36. [doi: 10.1177/1049732316658267] [Medline: 27469975]
23. Chagas BA, Pagano AS, Prates RO, Praes EC, Ferreguetti K, Vaz H, et al. Evaluating user experience with a chatbot designed as a public health response to the COVID-19 pandemic in Brazil: mixed methods study. JMIR Hum Factors 2023 Apr 04;10:e43135 [FREE Full text] [doi: 10.2196/43135] [Medline: 36634267]
24. Newton AS, March S, Gehring ND, Rowe AK, Radomski AD. Establishing a working definition of user experience for eHealth interventions of self-reported user experience measures with eHealth researchers and adolescents: scoping review. J Med Internet Res 2021 Oct 26;23(12):e25012 [FREE Full text] [doi: 10.2196/25012] [Medline: 34860671]
25. Abrahamson K, Hass Z, Arling G. Shall I stay or shall I go? The choice to remain in the nursing home among residents with high potential for discharge. J Appl Gerontol 2020 Oct 26;39(8):863-870. [doi: 10.1177/0733464818807818] [Medline: 30366510]
26. Cordes T, Bischoff LL, Schoene D, Schott N, Voelcker-Rehage C, Meixner C, et al. A multicomponent exercise intervention to improve physical functioning, cognition and psychosocial well-being in elderly nursing home residents: a study protocol of a randomized controlled trial in the PROCARE (prevention and occupational health in long-term care) project. BMC Geriatr 2019 Dec 23;19(1):369 [FREE Full text] [doi: 10.1186/s12877-019-1386-6] [Medline: 31870314]
27. Poncumhak P, Amput P, Sangkarit N, Promsrisuk T, Srithawong A. Predictive ability of the 2-minute step test for functional fitness in older individuals with hypertension. Ann Geriatr Med Res 2023 Jul 28;27(3):228-234 [FREE Full text] [doi: 10.4235/agmr.23.0070] [Medline: 37500190]
28. Zhang C, Dong X, Ding M, Chen X, Shan X, Ouyang H, et al. Executive control, alerting, updating, and falls in cognitively healthy older adults. Gerontology 2020 Aug 25;66(5):494-505. [doi: 10.1159/000509288] [Medline: 32841943]
29. Costenoble A, Knoop V, Vermeiren S, Vella RA, Debain A, Rossi G, et al. A comprehensive overview of activities of daily living in existing frailty instruments: a systematic literature search. Gerontologist 2021 Apr 03;61(3):e12-e22 [FREE Full text] [doi: 10.1093/geront/gnz147] [Medline: 31872238]
30. Liu-Ambrose T, Davis JC, Best JR, Dian L, Madden K, Cook W, et al. Effect of a home-based exercise program on subsequent falls among community-dwelling high-risk older adults after a fall: a randomized clinical trial. JAMA 2019 Jun 04;321(21):2092-2100 [FREE Full text] [doi: 10.1001/jama.2019.5795] [Medline: 31162569]
31. Gazibara T, Kurtagic I, Kisic-Tepavcevic D, Nurkovic S, Kovacevic N, Gazibara T, et al. Falls, risk factors and fear of falling among persons older than 65 years of age. Psychogeriatrics 2017 Jan 27;17(4):215-223. [doi: 10.1111/psyg.12217] [Medline: 28130862]

Abbreviations

ADL: activities of daily living
GLF: ground-level fall
HPD: head protection device
LOS: length of stay
UX: user experience

XSL•FO
RenderX

Edited by A Kushniruk; submitted 26.11.23; peer-reviewed by DD Larson MD, S Li, D Low; comments to author 31.01.24; revised version received 08.02.24; accepted 18.02.24; published 19.03.24.

Please cite as:
Haag S, Kepros J
Head Protection Device for Individuals at Risk for Head Injury due to Ground-Level Falls: Single Trauma Center User Experience Investigation
JMIR Hum Factors 2024;11:e54854
URL: https://humanfactors.jmir.org/2024/1/e54854
doi:10.2196/54854
PMID:38502170

©Susan Haag, John Kepros. Originally published in JMIR Human Factors (https://humanfactors.jmir.org), 19.03.2024. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on https://humanfactors.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

A Digital Health Intervention Platform (Active and Independent Management System) to Enhance the Rehabilitation Experience for Orthopedic Joint Replacement Patients: Usability Evaluation Study

Petros Papadopoulos[1*], BSc, PhD; Mario Soflano[2*], BSc, PhD; Thomas Connolly[3*], BSc, PhD

[1]University of Strathclyde, Glasgow, United Kingdom
[2]Glasgow Caledonian University, Glasgow, United Kingdom
[3]DS Partnership, Glasgow, United Kingdom
[*]all authors contributed equally

Corresponding Author:
Petros Papadopoulos, BSc, PhD
University of Strathclyde
26 Richmond Street
Glasgow, G1 1XH
United Kingdom
Phone: 44 7920756076
Email: petros.papadopoulos@gmail.com

Abstract

Background: Optimal rehabilitation programs for orthopedic joint replacement patients ensure faster return to function, earlier discharge from hospital, and improved patient satisfaction. Digital health interventions show promise as a supporting tool for re-enablement.

Objective: The main goal of this mixed methods study was to examine the usability of the AIMS platform from the perspectives of both patients and clinicians. The aim of this study was to evaluate a re-enablement platform that we have developed that uses a holistic systems approach to address the *de-enablement* that occurs in hospitalized inpatients, with the older adult population most at risk. The Active and Independent Management System (AIMS) platform is anticipated to deliver improved patient participation in recovery and self-management through education and the ability to track rehabilitation progression in hospital and after patient discharge.

Methods: Two well-known instruments were used to measure usability: the System Usability Scale (SUS) with 10 items and, for finer granularity, the User Experience Questionnaire (UEQ) with 26 items. In all, 26 physiotherapists and health care professionals evaluated the AIMS clinical portal; and 44 patients in hospital for total knee replacement, total hip replacement, or dynamic hip screw implant evaluated the AIMS app.

Results: For the AIMS clinical portal, the mean SUS score obtained was 82.88 (SD 13.07, median 86.25), which would be considered *good/excellent* according to a validated adjective rating scale. For the UEQ, the means of the normalized scores (range −3 to +3) were as follows: attractiveness=2.683 (SD 0.100), perspicuity=2.775 (SD 0.150), efficiency=2.775 (SD 0.130), dependability=2.300 (SD 0.080), stimulation=1.950 (SD 0.120), and novelty=1.625 (SD 0.090). All dimensions were thus classed as *excellent* against the benchmarks, confirming the results from the SUS questionnaire. For the AIMS app, the mean SUS score obtained was 74.41 (SD 10.26), with a median of 77.50, which would be considered *good* according to the aforementioned adjective rating scale. For the UEQ, the means of the normalized scores were as follows: attractiveness=2.733 (SD 0.070), perspicuity=2.900 (SD 0.060), efficiency=2.800 (SD 0.090), dependability=2.425 (SD 0.060), stimulation=2.200 (SD 0.010), and novelty=1.450 (0.260). All dimensions were thus classed as *excellent* against the benchmarks (with the exception of novelty, which was classed as *good*), providing slightly better results than the SUS questionnaire.

Conclusions: The study has shown that both the AIMS clinical portal and the AIMS app have *good* to *excellent* usability scores, and the platform provides a solid foundation for the next phase of research, which will involve evaluating the effectiveness of the platform in improving patient outcomes after total knee replacement, total hip replacement, or dynamic hip screw.

XSL•FO
RenderX

(JMIR Hum Factors 2024;11:e50430) doi:10.2196/50430

KEYWORDS

mobile health; mHealth; digital health intervention; total knee replacement; TKR; total hip replacement; THR; dynamic hip screw; DHS; rehabilitation; usability; mobile phone

Introduction

Background

According to the World Health Organization's 2019 Global Burden of Disease study, approximately 1.71 billion people globally experience musculoskeletal conditions. Low back pain is the most common condition, affecting an estimated 568 million people [1]. In the United Kingdom, it has been estimated that musculoskeletal conditions affect >20 million people, approximately a third of the population [2]. Musculoskeletal conditions are the second greatest contributor to disability worldwide and is a significant burden to the individual and society [3]. It is expected that the impact of musculoskeletal conditions on the health service and on society will continue to rise as life expectancy increases [4]. Many different approaches have been explored to reduce this burden, including medical interventions, work-related approaches (reducing stress at work as well as improving health and safety regulations), social education (improving awareness of exercise and healthy eating), and the use of technology.

Musculoskeletal conditions comprise >150 different disorders, diseases, and syndromes that affect bones, joints, muscles, the spine, and soft tissues [3]. While some conditions are short lived, such as sprains and fractures, others can be lifelong conditions requiring ongoing treatment. Pain is a common symptom of musculoskeletal conditions. Back and neck pain, osteoarthritis, rheumatoid arthritis, and fractures are among the most disabling conditions and can be a significant barrier to healthy aging [5]. Musculoskeletal conditions can be classified by the body part affected (eg, knee pain and shoulder pain), whether the condition is noninflammatory (such as osteoarthritis) or inflammatory (such as rheumatoid arthritis), and whether the condition is restricted to the musculoskeletal system or more widespread (such as systemic lupus erythematosus) [4]. To compound matters, musculoskeletal issues tend to be associated with other diseases, such as heart or respiratory disease and stroke, and lead to an increase in disabilities and deaths [6-8]. It has been estimated that musculoskeletal conditions account for up to 21% of annual general practitioner consultations across England [9], and health service costs from inability to work and sickness absence in the United Kingdom are approximately £100 billion (US $125 billion) annually [10]. It is important to find solutions that will help reduce the significant burdens on the individual, society, the economy, and the health service. While many solutions will be of a medical nature, technology has a significant part to play in easing the burdens. In the next subsection, we discuss some digital health interventions (DHIs) for musculoskeletal conditions.

A number of different terminologies exist in the health domain for software solutions generally. The terms eHealth and mobile health (mHealth) have been used for a number of years. More recently, the more encompassing term *digital health* has been introduced. This is defined as "encompassing eHealth [which includes mHealth] as well as developing areas such as the use of advanced computing sciences (in the fields of 'big data,' genomics and artificial intelligence, for example)" [11]. Examples of digital health solutions include primary and secondary care IT systems; patient portals that provide secure web-based access to a range of health services, such as My Diabetes My Way and PatientView [12]; personal health data stores such as Mydex [13]; telehealth systems such as Attend Anywhere and Near Me [14]; and health-related mobile apps. It is believed that these systems can benefit health care delivery by improving different outcomes, such as effectiveness, efficiency, accessibility, safety, and personalization [15]. There has been a growing public interest in DHIs because they can allow individuals to monitor, manage, and improve their health and quality of life in a more personalized way, potentially more cost-effectively, and at a time that suits them [16-18].

Optimal rehabilitation programs for orthopedic joint replacement patients ensure faster return to function, earlier discharge from hospital, and improved patient satisfaction [4,19-21] as well as prevent further deconditioning [22]. The aim of this study was to evaluate the usability of a re-enablement platform called Active and Independent Management System (AIMS) that was developed to address the *de-enablement* that occurs in hospitalized inpatients for one of the groups considered to be most at risk, that is, older adults. The platform is capable of delivering digital rehabilitation plans and tracking the progression of the plans in real time; in addition, it can be used both in hospital and at home after a patient is discharged. The rationale for using such a system is to help reduce the time spent in hospital and improve patient satisfaction through self-management.

Re-Enablement DHIs

This subsection examines some recent literature related to the use of DHIs for total knee replacement (TKR) or total knee arthroplasty (TKA) and total hip replacement (THR) or total hip arthroplasty (THA). Hussain et al [23] developed a TKR platform comprising a mobile phone app, a wrist-worn activity tracker, and a clinical web portal. The purpose-built iOS and Android apps included weekly psychoeducation sessions and tasks that were delivered by a program guide via text and voice recordings. By obtaining the data from the tracker and the app, the clinician could monitor patient progress and the configured physiotherapy programs, while the patient care team could review the progress and the designated programs using the web portal. Physiotherapy programs were mostly from a library of videos created for TKR rehabilitation, which were made available in the app once set by the clinician. The authors planned to conduct a 13-month multisite unblinded randomized controlled trial in which participants were assigned to 1 of 2 study groups [23]. The participants for the experiment were patients who underwent TKR, and the study included an active

XSL•FO
RenderX

intervention period from the time the patients were scheduled for surgery (approximately 4 weeks before surgery) to 12 weeks after the surgery, followed by a 40-week free-living period until 1 year after surgery.

Timmers et al [24] investigated the effect of a mobile app for day-to-day postoperative care education on TKR patients regarding the level of pain compared to those who only received standard information about their recovery through the app. The study involved 114 patients in the intervention group and 99 patients in the control group. In the intervention group, 93 patients downloaded and used the app. The results showed that, in comparison with standard patient education, the active education and coaching of patients on a day-to-day basis via the app in the 4 weeks after TKR resulted in a significant decrease, among other things, in the patients' levels of pain and a significant improvement in patients' physical functioning and quality of life, as well as their ability to perform physiotherapy exercises and activities of daily self-care.

Van Dijk-Huisman et al [25] developed a mobile app to prevent the negative effects of inactivity in hospital. The app supported objective activity monitoring, gave patients a view of their recovery progress, and offered a customized exercise program. The aim of the study was to investigate the potential of the app to enhance physical activity levels and functional recovery after orthopedic surgery discharge. In all, 97 patients undergoing TKA and THA were recruited for the evaluation. The control group (n=64) received standard physiotherapy, while the intervention group (n=33) used the mobile app in addition to physiotherapy. The time spent in active and functional recovery on postoperative day 1 (POD1) was measured. The app use, corrected for age, resulted in patients standing and walking on POD1 for an average increase of 28.43 (95% CI 5.55-51.32) minutes. The odds of achieving functional recovery on POD1 were 3.08 times higher (95% CI 1.14-8.31) with the use of the mobile app. The authors concluded that a mobile app combined with an accelerometer demonstrated the potential to enhance patients' activity levels and functional recovery during their hospital stay [25].

Wijnen et al [26] investigated the effectiveness of a home-based rehabilitation program using a tablet app and remote coaching for patients after THA. Existing data from 2 studies were combined: patients from a single-arm intervention study were matched with the historical controls from an observational study. Patients aged 18 to 65 years who had undergone THA were included. The intervention group had a 12-week home-based rehabilitation program with instructional videos on a tablet device and remote coaching. Patients were asked to perform strengthening and walking exercises at least 5 days a week. The intervention group was compared with a control group that included patients who received usual care. Effectiveness was measured at 4 points (preoperatively and 4 weeks, 12 weeks, and 6 months postoperatively) by means of functional tests and self-reported questionnaires. The intervention group performed functional tests significantly faster at 12 weeks and 6 months postoperatively and also scored significantly higher on the subscales *function in sport and recreational activities* and *hip-related quality of life* of the Hip Disability and Osteoarthritis Outcome Questionnaire, as well as on the subscale *physical role limitations* of the Short Form Health Survey-36 at 12 weeks and 6 months postoperatively. Large effect sizes were found on functional tests at 12 weeks and 6 months, endorsed by effect sizes on the self-reported outcomes. The authors concluded that the results demonstrated larger effects in the intervention group than in the historical controls, indicating that a home-based rehabilitation program using a mobile app after THA can be more effective than usual care [26].

Bell et al [27] ran a controlled pilot study for TKR patients, investigating the feasibility and effectiveness of interACTION, a remote (wearable) rehabilitation monitoring platform developed for use by patients after TKR. The InterACTION platform has portable motion sensors placed on either side of a joint to collect joint orientation data using a custom mobile app and then send the data to the clinician's web-based portal. The mobile app also contains 30 knee-specific home exercises for TKR rehabilitation that the physical therapist can personalize remotely through a web-based clinical portal. The study compared 2 groups: 19 patients who used the interACTION platform and a control group with 19 patients who used standard postoperative outpatient rehabilitation with a physical therapist (2-3 sessions per week over a maximum of 10 weeks), supplemented with a home exercise program. The primary outcome measured was value, operationally defined as the change in the activities of daily living scale of the Knee Outcome Survey at 10 weeks divided by the total cost of rehabilitation (determined from the total number of physical therapy sessions and the billable charges for each session during the 10 weeks the patients were enrolled in the study). In terms of this measure, no statistical differences were found between the groups. The study showed relatively low and not significant differences between the groups in terms of attrition rates, indicating that both interventions were acceptable. There was a small decrease in clinic visits by patients in the interACTION group, and all patients and physical therapists in the group indicated that they would use the system again.

Bäcker et al [28] developed a mobile app with a GenuSport sensor that allows isokinetic exercises to improve postoperative quadriceps weakness and knee motion. The sensor was placed underneath the patient's knee, and gamified exercise routines were presented through the app consisting of two exercises: (1) *high striker game*, where the patient has to push the knee onto the sensor for 5 seconds; and (2) *flight simulator*, where the player is supposed to keep the knee in the air for 100 seconds. The authors carried out a randomized controlled trial with a 2-year follow-up to evaluate the effectiveness of the app-based rehabilitation for patients after TKA [28]. In all, 35 patients completed the study and were randomly assigned to 2 groups: 20 patients received the app-based exercise program, and 15 patients were included in the control group. Patients in the app group used an external device to measure knee range of motion starting on the day of surgery, whereas patients in the control group underwent regular physiotherapy. Functional outcome scores using the Knee Injury and Osteoarthritis Outcome Score, the Knee Society Scoring System, and a visual analog scale for pain were analyzed. The results showed that, in the short term, the app group performed significantly better than the control group when taking a 10-minute walk, with less pain. In the

XSL•FO
RenderX

longer term, the app group also performed significantly better, with higher Knee Society Scoring System scores as well as requiring fewer painkillers. In addition, the app group participants were more likely to participate in sports.

Colomina et al [29] developed an mHealth system for older patients with complex chronic conditions undergoing elective THA or TKA. The mHealth system formed part of the Personalized Connected Care for Complex Chronic Patients platform, which contained a web-based smart adaptive case management system for health care professionals that seamlessly integrated with a patient self-management mHealth system that supported communication between health care professionals and patients. The authors assessed the effectiveness and cost-effectiveness of implementing an mHealth-enabled integrated care (IC) model for patients with complex chronic conditions undergoing TKA or THA versus usual care [29]. A prospective pragmatic 2-arm parallel implementation trial was conducted in the rural region of Lleida in Catalonia, Spain, for 3 months. A total of 29 patients with complex chronic conditions undergoing TKA or THA and their caregivers received the IC program, while 30 patients with statistically comparable baseline characteristics, such as age, sex, and type of arthroplasty, were recruited for the usual care group. The results suggested that both treatment models significantly improved the physical and mental health status of the patients; however, IC significantly reduced the number of unplanned visits related to the surgery procedure and consequently significantly lowered the patients' expenses.

Rian et al [30] presented a web tool called Eir for symptom registration at home after knee arthroplasty. Given that the system was previously used in cancer care, a separate patient module was designed for patient-reported postoperative symptom assessment and medication registration after fast-track TKA that consisted of measurements of pain and side effects, as well as detailed registration of the use of analgesic drugs. The authors conducted a usability and feasibility study using a randomized controlled trial involving 134 participants [30]. The tool's usability was assessed with the use of the System Usability Scale (SUS) by 119 of the 134 participants, while the feasibility data were collected qualitatively. The results showed that 70% of the participants managed to use the tool at home without any technical support, although they indicated technical challenges related to the log-in procedure or internet access. The usability was rated high, with a mean SUS score of 89.6 (median 92.5; range 22.5-100).

Two literature reviews assessing the use of app-based rehabilitation for TKA or THA were conducted recently [31,32]. Bäcker et al [31] examined the functional outcomes of app-based rehabilitation of patients after TKA or THA. The review identified 420 entries from MEDLINE or PubMed and Google databases, but only 9 publications met the inclusion criteria, covering 518 patients in the intervention groups and 549 patients in the control groups. Five studies used app-based exercise instructions delivered via a mobile device, and 4 studies used a sensor or motion tracker. The average follow-up was 9.5 (SD 8.1; range 3-23.4) months. Overall, significantly lower activity visual analog scale values were observed for the interventional groups in the short term (P=.002). There were no other significant differences observed between the 2 groups. The study found that there were significant short-term improvements in the mobile app group. The authors concluded that mobile apps provide an alternative to in-person sessions that may improve access to physical activity for patients after TKA or THA, and, in combination with a Bluetooth-enabled sensor for isometric exercises, patients can additionally receive real-time feedback after TKA or THA [31].

Constantinescu et al [32] conducted a systematic literature review on the use of commercially available smartphone apps and wearable devices to assist rehabilitation interventions after TKA from the PubMed, Cochrane Library, MEDLINE, and Web of Science databases. Of the 60 full-text studies identified (published between January 2020 and September 2021), a total of 15 met the inclusion criteria, of which 4 studies used smartphone apps, 7 used wearable devices, and 4 used both to monitor physical activity and patient status after TKA. In terms of primary outcomes, 3 studies examined device accuracy, 3 recovery prediction, 2 functional recovery, 2 physical activity promotion, 2 patient compliance, 2 pain control, and 1 study examined health care use. The authors concluded that commercially available apps and wearable devices can capably monitor physical activity and improve patient engagement after TKA, making them approaches that support or replace traditional rehabilitation programs [32]. Using different strategies in interventions, such as setting step goals, using app-based patient engagement platforms, and establishing patient-specific benchmarks for recovery, can enhance the effectiveness of the treatment.

The AIMS Platform

It is well established that musculoskeletal conditions contribute to a large number of disabilities worldwide, and the projections show that this number will continue to rise. Initiatives are ongoing to combat this problem proactively (eg, reducing stress, improving health and safety regulations, exercising, and healthy eating). Furthermore, reactive approaches of optimal prehabilitation and rehabilitation programs are also undergoing development to optimize operations and ensure the best use of available resources while improving patient satisfaction. The aim of this study was to evaluate a rehabilitation platform in an effort to combat the lack of enablement in hospitalized older adults considered more vulnerable. The platform is capable of delivering digital rehabilitation plans and tracking the progression of these plans in real time; in addition, it can be used both in hospital and at home after patient discharge. The rationale for using such a system is to help reduce time spent in hospital and improve patient satisfaction through self-management.

The AIMS platform helps manage patients' rehabilitation programs. Each patient is registered by a clinician at the beginning of their patient pathway, and the system collects certain relevant information about the patient as they move through their journey. Rehabilitation clinicians use this platform to create, monitor, and adjust a patient's rehabilitation package as and when required. A team consisting of stakeholders is assigned to each patient and is responsible for the delivery of the program. A library of physiotherapy exercises and

XSL•FO
RenderX

educational videos (eg, the use of a walking aid or how to apply a patient's splint) have been recorded and uploaded into the system (Figure 1). Each staff member can attach a series of video exercises specific to the patient's needs that can help the re-enablement process.

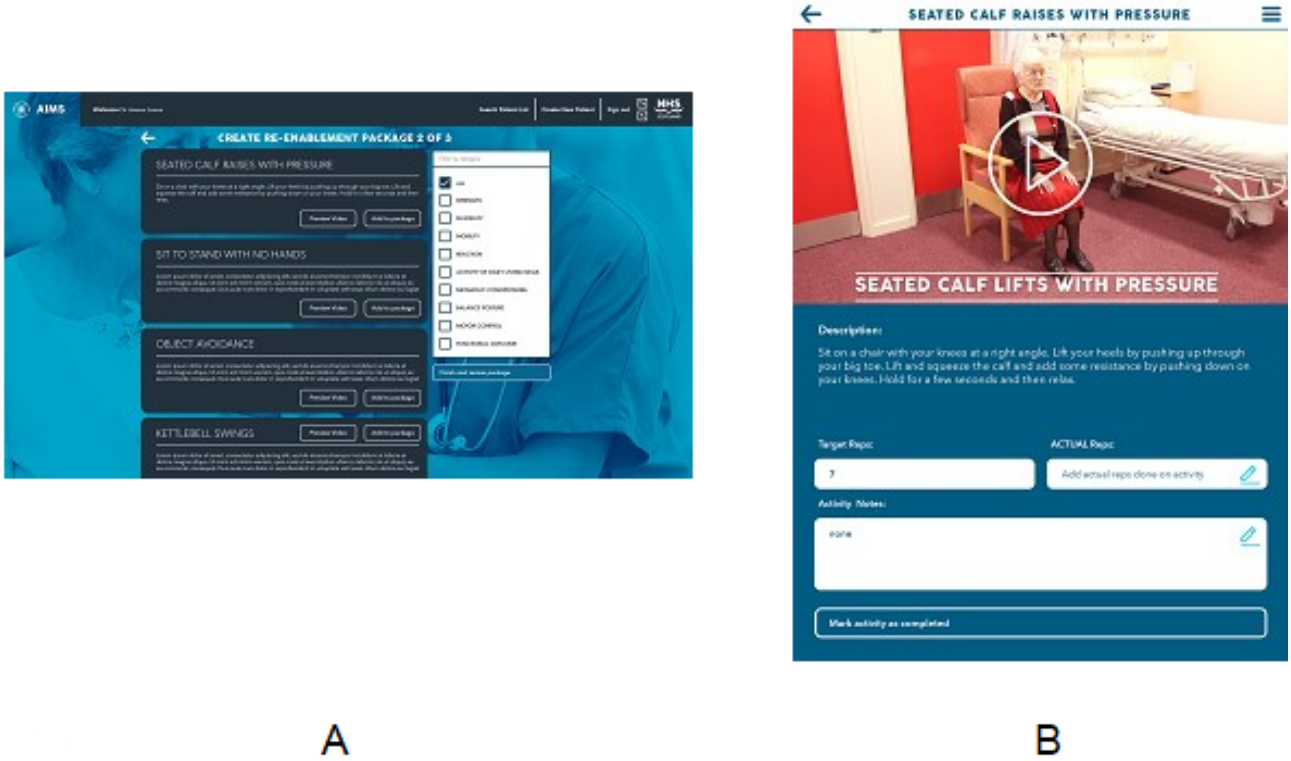
The platform consists of 2 components: a web content management system used by clinicians to create rehabilitation plans for postoperative patients and a mobile app designed to deliver these rehabilitation plans to the patients with a series of exercises to be completed by them.

The clinician starts by creating a user account for a patient (all patient information is anonymized with a random unique ID number that is later used to gain access to the rehabilitation plan through the app; Figure 1). The clinician then sets up a specific rehabilitation plan for this patient according to their needs. The clinician can search from all available exercises by using general search terms to filter what is available and also preview the associated video to make sure the appropriate plan is created (Figure 1A). The clinician must then determine the number of repetitions for each exercise and the frequency at which they need to be performed each day, usually 4 sessions a day. After the plan is complete, it becomes active on the patient's device, and the clinician then demonstrates to the patient how to use the app. This enables the clinician to monitor a patient and their progress each time they complete an exercise. There is also capability for the patient to comment on any particular issues with any of the exercises, and the clinician will be able to view the comments using the content management system and modify the plan accordingly. A typical example of a rehabilitation plan for a patient consists of some simple but very effective exercises (eg, heel slides, knee extensions, and knee flexions). The patient would be asked to perform 10 repetitions of all exercises in 4 daily sessions. The rehabilitation process starts after the operation for as long as the patient remains in hospital, and there are physiotherapists available to offer assistance during the patient rehabilitation process; the app does not prompt patients to complete their daily rehabilitation plan because these sessions are already scheduled in the hospital ward. It is up to the patient to continue using the app for rehabilitation after hospital discharge (the app is available for free download from app stores).

The patient uses a tablet device provided by the hospital to gain access to the AIMS mobile app and work on their rehabilitation plan. Each user is given a random ID number generated by the clinician that is required to log in to the app; no password is required because all information is anonymized. After this, the user can use the app and work on their specific rehabilitation plan and set of exercises and also view their daily plan progress. The patient is provided a textual description of the exercise and a video with audio explaining how it should be performed and how many repetitions should be performed (Figure 1B). At the bottom of the page featuring each exercise, feedback can be provided on how many repetitions were achieved as well as any comments if there were any issues when performing the exercise.

Typically, the app would be used by a member of the staff or a member of the family during visiting hours to help the patient with their exercises by encouraging them or participating with them and achieving successful completion of the rehabilitation plan.

Figure 1. (A) The Create Rehabilitation Re-Enablement Package Screen, and (B) the patient exercise screen.



XSL•FO
RenderX

Aims of the Study

The aim of this study was to investigate the usability of the AIMS platform from the perspectives of both clinicians and patients. Two well-known instruments were used to measure usability: the SUS [33] with 10 items and, for finer granularity, the User Experience Questionnaire (UEQ) [34] with 26 items.

The evaluation aims to answer the following 2 research questions (RQs):

- RQ1: does the AIMS clinical portal provide a solution that could be usable by clinicians?
- RQ2: does the AIMS app provide a solution that could be usable by patients?

Methods

Overview

The World Health Organization defines evaluation as "the systematic and objective assessment of an ongoing or completed project [with the aim of determining] the relevance and fulfilment of objectives, development efficiency, effectiveness, impact and sustainability" [35]; and the guide for monitoring and evaluating DHIs outlines 7 stages of DHI maturity, ranging from preprototype to full deployment. This project is considered to be at the prototype stage of maturity, which would include usability testing. Ways to improve the system would also be investigated.

Usability is recognized as a significant quality indicator that determines the success of software applications [36-39]. Johnson et al [40] defines three main approaches to evaluate usability: (1) user based (a sample of prospective users use the system), (2) expert based (≥1 usability or human-computer interaction experts evaluate the system), and (3) model based (formal methods are used to predict user performance). Our health board members were keen on using the user-based approach to evaluate the DHI; hence, this approach was chosen.

Many validated usability instruments have been proposed in the literature with varying numbers of questions. In this study, we used 2 well-known validated instruments: the SUS and the UEQ. The SUS [33] consists of 10 statements (5 positive and 5 negative) that the users rate on a scale ranging from 1=*strongly disagree* to 5=*strongly agree*. The questionnaire alternates between positive and negative statements to avoid random answers. The aggregated score out of 100 can be compared with the average SUS benchmark score of 68.0. To represent SUS scores, Bangor et al [41] defined a 7-point adjective rating scale: *best imaginable*, *excellent*, *good*, *OK*, *poor*, *awful*, and *worst imaginable*.

The UEQ assesses the extent to which (1) the product meets expectations and (2) a product can be compared with other systems using a published benchmark. Schrepp et al [42] developed an adjective rating scale for benchmarking, and a mean score of >1.75 would be considered in the 10% best results. While the UEQ provides finer detail than the SUS, it was felt that asking busy clinicians to rate 26 statements may result in a smaller number of responses compared to asking them to rate 10 SUS statements; therefore, it was decided to use the SUS with all participants and the UEQ with a small number of participants.

Some qualitative information was also gathered using open-ended questions to gain a deeper understanding of participants' views of the AIMS platform.

Ethical Considerations

Ethics approval for this study was obtained from Hairmyres University Hospital, Lanarkshire. One of the conditions of approval was that all personal information from the study should be removed and that patient information should be kept private and safe (Data Protection Impact Assessment Questionnaire for Active Independent Mobility System [AIMS] Pilot Study Hairmyres University Hospital, Lanarkshire; May 29, 2019). All participants provided consent before participating in the study.

Participants

In all, 26 physiotherapists and health care professionals volunteered to evaluate the AIMS clinical portal; and 44 patients in hospital for TKR, THR, or dynamic hip screw (DHS) agreed to participate in evaluating the AIMS app. The study was carried out on May 5, 2019, or November 14, 2019.

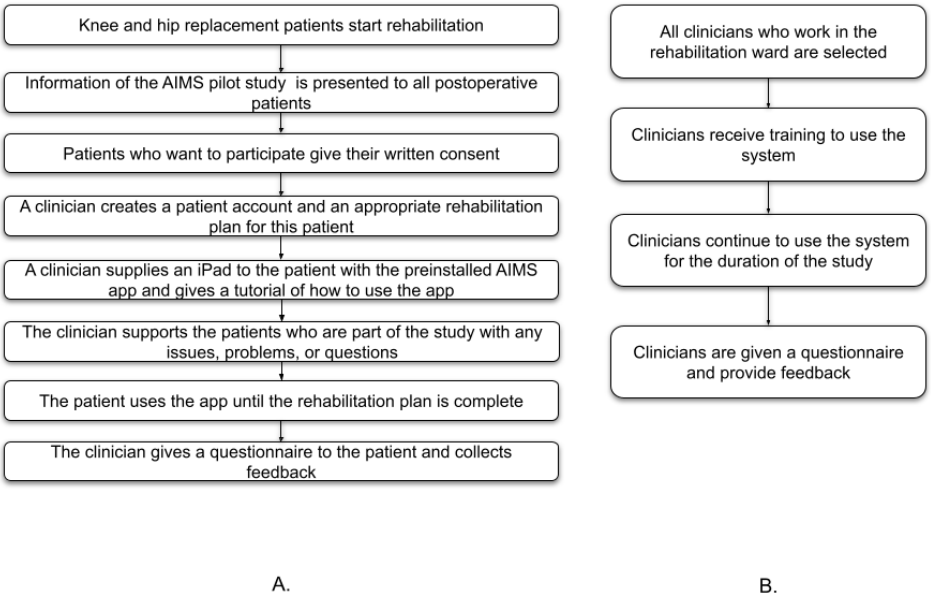
Test Protocol

The 6-month-long study was undertaken in the rehabilitation ward in an Hairmyres University Hospital, Lanarkshire, that specializes in TKR or THR surgery. Only clinicians had access to the patients during their stay in the hospital, and a member of the team of clinicians (lead study clinician) had oversight over recruiting and running the experiment. Technical support was provided by the study team to all clinicians during the rehabilitation sessions in the hospital, but this was not in the rehabilitation ward. Questionnaires were given to the participants in a paper-based form, which they were asked to complete and hand back to the lead study clinician toward the end of their rehabilitation stay. The study was conducted using 10 hospital-supplied second-generation iPad Air 2 devices running iOS 10.3 with a 9.7-inch display in portrait orientation (refer to the patient test protocol presented in Figure 2A).

Clinicians of the rehabilitation team in the hospital were all given training on how to use the portal to create rehabilitation packages for patients and how they would look in the app. They were also involved in the development and design process with focus groups and early prototyping, which enabled most of them to develop a good understanding of the AIMS platform. As the study was taking place alongside patients who were not part of the study, everyone had to be able to help the patients, which is why they were all trained to use the system. Clinicians were given a questionnaire to complete after they had used the platform a few times (refer to the clinician test protocol presented in Figure 2B).

XSL•FO
RenderX

Figure 2. (A) Patient test protocol. (B) Clinician test protocol. AIMS: Active and Independent Management System.



Results

For the SUS, the analysis was carried out using Excel (Microsoft Corp); and for the UEQ, the analysis was carried out using the standard UEQ spreadsheet.

AIMS Clinical Portal: SUS Results

All 26 participants completed the SUS questionnaire (100% response rate). The mean SUS score obtained was 82.88 (SD 13.07), with a median of 86.25. This score would be considered *good/excellent* according to the adjective rating scale developed by Bangor et al [41]. A breakdown of the participants' answers to the SUS questions regarding the AIMS clinical portal is provided in Table 1.

Table 1. Participants' answers to the System Usability Scale (SUS) questions for the Active and Independent Management System clinical portal (n=26).

Statements	Participants agreeing, n (%)	Participants disagreeing, n (%)	SUS scores, mean (SD)
Positive statements			
I think that I would like to use this system frequently	23 (88)	2 (8)	4.00 (0.76)
I thought the system was easy to use	24 (92)	1 (4)	4.50 (0.77)
I found the various functions in this system were well integrated	22 (85)	2 (8)	4.50 (0.96)
I would imagine that most people would learn to use this system very quickly	23 (88)	1 (4)	4.19 (0.65)
I felt very confident using this system	24 (92)	1 (4)	4.38 (1.42)
Negative statements			
I found the system unnecessarily complex	2 (8)	23 (88)	2.04 (0.73)
I think that I would need the support of a technical person to be able to use this system	1 (4)	24 (92)	1.50 (0.77)
I thought there was too much inconsistency in this system	0 (0)	23 (88)	1.62 (0.86)
I found the system very cumbersome to use	3 (88)	21 (81)	1.73 (1.05)
I needed to learn a lot of things before I could get going with this system	0 (0)	24 (92)	1.42 (0.65)

AIMS Clinical Portal: UEQ Results

Invitations were sent to 12 (46%) of the 26 participants. Of these 12 participants, 10 (83%) completed the UEQ questionnaire. The means of the normalized scores (range −3 to +3) for the AIMS clinical portal were as follows: attractiveness=2.683 (SD 0.100), perspicuity=2.775 (SD 0.150), efficiency=2.775 (SD 0.130), dependability=2.300 (SD 0.080), stimulation=1.950 (SD 0.120), and novelty=1.625 (SD 0.090). Figure 3 shows the bar chart of the results for the AIMS clinical

XSL•FO
RenderX

portal against the benchmarks, showing all dimensions classed as *excellent* and confirming the results from the SUS questionnaire. Table 2 provides the mean (SD) and variance of the normalized values for the items in the UEQ questionnaire for the AIMS clinical portal. In most cases, the values are very encouraging, with the exception of *conservative* and *innovative*, although this is still rated *good*. Figure 4 shows the bar chart of the data grouped into the 6 UEQ dimensions.

Figure 3. Bar chart of the Active and Independent Management System clinical portal User Experience Questionnaire results against the benchmarks.

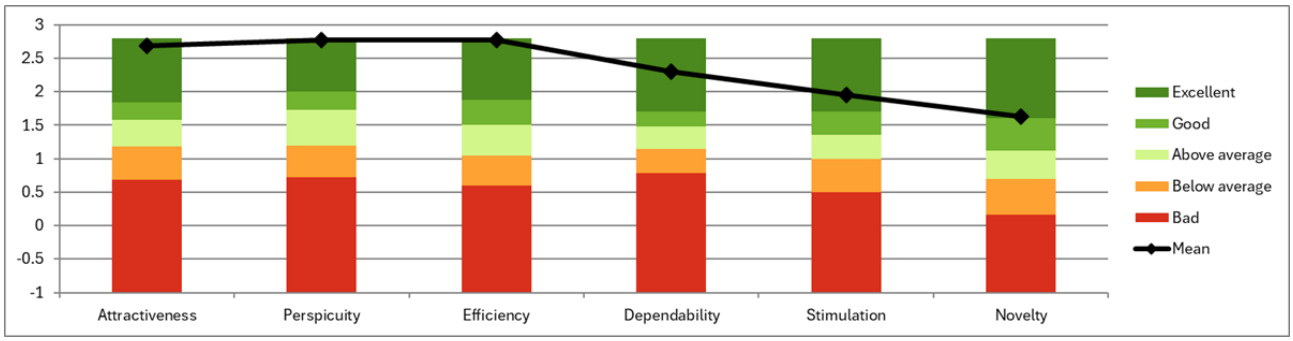
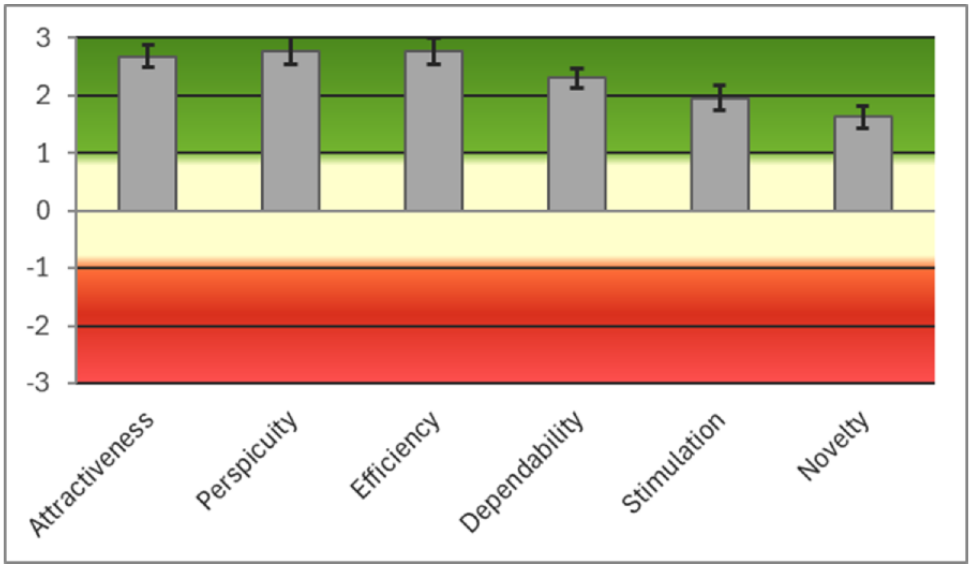


Table 2. Mean (SD) and variance of the normalized values for the items in the User Experience Questionnaire (UEQ) for the Active and Independent Management System clinical portal (n=10).

Scale	Left anchor of the scale	Right anchor of the scale	UEQ scores, mean (SD)	Variance
Attractiveness	Annoying	Enjoyable	1.9 (0.3)	0.1
Perspicuity	Not understandable	Understandable	2.6 (0.7)	0.5
Novelty	Creative	Dull	1.9 (0.3)	0.1
Perspicuity	Easy to learn	Difficult to learn	2.9 (0.3)	0.1
Stimulation	Valuable	Inferior	2.2 (0.6)	0.4
Stimulation	Boring	Exciting	1.5 (0.5)	0.3
Stimulation	Not interesting	Interesting	2.1 (0.6)	0.3
Dependability	Unpredictable	Predictable	1.8 (0.4)	0.2
Efficiency	Fast	Slow	2.8 (0.4)	0.2
Novelty	Inventive	Conventional	1.9 (0.3)	0.1
Dependability	Obstructive	Supportive	2.2 (0.6)	0.4
Attractiveness	Good	Bad	2.8 (0.4)	0.2
Perspicuity	Complicated	Easy	2.7 (0.5)	0.2
Attractiveness	Unlikable	Pleasing	2.7 (0.5)	0.2
Novelty	Usual	Leading edge	1.2 (0.4)	0.2
Attractiveness	Unpleasant	Pleasant	2.9 (0.3)	0.1
Dependability	Secure	Not secure	2.5 (0.5)	0.3
Stimulation	Motivating	Demotivating	2.0 (0.5)	0.2
Dependability	Meets expectations	Does not meet expectations	2.7 (0.5)	0.2
Efficiency	Inefficient	Efficient	2.8 (0.4)	0.2
Perspicuity	Clear	Confusing	2.9 (0.3)	0.1
Efficiency	Impractical	Practical	2.7 (0.5)	0.2
Efficiency	Organized	Cluttered	2.8 (0.4)	0.2
Attractiveness	Attractive	Unattractive	2.9 (0.3)	0.1
Attractiveness	Friendly	Unfriendly	2.9 (0.3)	0.1
Novelty	Conservative	Innovative	1.5 (0.5)	0.3

XSL•FO
RenderX

Figure 4. Bar chart of the Active and Independent Management System clinical portal data grouped into the 6 User Experience Questionnaire dimensions.



AIMS Clinical Portal: Qualitative Feedback

To gain further insight into how users perceived the AIMS clinical portal, 3 additional questions were asked (refer to the following subsections).

Q1: What Do You Think Are the Advantages of This Portal?

Of the 26 participants, 20 (77%) answered this question. All clinicians (20/20, 100%) who answered the question thought that providing customized exercise videos after an operation was very useful, particularly for patients being able to use the system at home; 15 (75%) of the 20 clinicians also suggested that receiving immediate feedback on how patients were coping with the exercise regime was very helpful and meant that the regime could be easily customized for each patient based on how they were coping, which was a key advantage. In addition, 60% (12/20) of the clinicians considered ease of use an advantage. Example comments were as follows:

> *Really liked the exercise videos for the patients; they were professionally produced and highly relevant for rehabilitation.* [Physiotherapist A]
>
> *I'm pleased to see that patients automatically receive some feedback on how they are progressing with the rehabilitation exercises.* [Physiotherapist B]

Q2: What Do You Think Are the Disadvantages of This Portal?

Of the 26 participants, 12 (46%) answered this question. Of these 12 clinicians, 7 (58%) thought that the integration of the portal with the current IT systems may be a challenge, 3 (25%) thought that getting the staff to agree to use the portal may be a possible issue, and 2 (17%) thought that some staff members would need training on how to use it. Example comments were as follows:

> *One big issue that will have to be addressed at some point is integrating the software with hospital systems, as we ultimately need to have the patient progress data in their EHR [electronic health record].* [Physiotherapist B]
>
> *While the current system was intuitive and easy to use, I wonder whether some training will need to be provided when the features to add further videos and provide more customised feedback are added.* [Physiotherapist C, an academic]

Q3: Would You Change Anything?

Of the 26 participants, 17 (65%) answered this question. Of these 17 clinicians, 12 (71%) suggested that the ability to create more self-help advice for patients would be useful, and 5 (29%) suggested that having a larger data bank of exercise regimes would be helpful. Example comments were as follows:

> *It would be very helpful if more self-help could be added to the app to reduce the dependency on the volume of information sheets we provide to patients.* [Physiotherapist D]
>
> *The current set of videos are very relevant and of a high quality; however, it would be beneficial to be able to have a wider selection of videos to be able [to] select from.* [Physiotherapist E]

AIMS App: SUS Results

The participants were selected during their first postoperative rehabilitation session (opportunistic recruitment). The recruitment of patients was carried out by a physiotherapist who would ask patients during their first session whether they were willing to participate in the study. The physiotherapist provided an information leaflet that explained what the study was about and how it could be used. All postoperative patients automatically qualified for the study; no one was excluded based on age, sex, or technical competency. The study did not collect any age- or sex-related information (a condition of the ethics approval for the study); therefore, it was not possible to provide information about patient demographics.

Of the 44 patients, 38 (86%) completed the SUS questionnaire. The mean SUS score obtained was 74.41 (SD 10.26), with a median of 77.50. This score would be considered *good* according

XSL•FO
RenderX

to the adjective rating scale developed by Bangor et al [41]. A breakdown of the participants' answers to the SUS questions for the AIMS app is provided in Table 3.

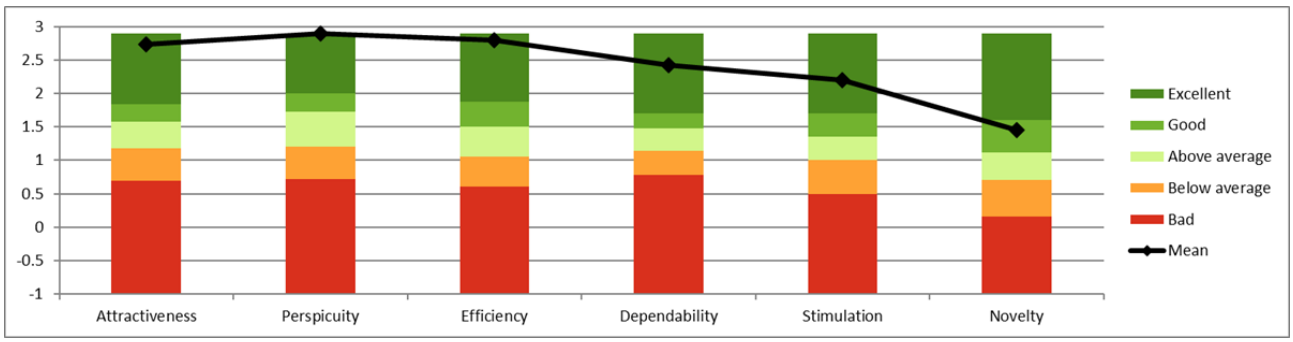
Table 3. Participants' answers to the System Usability Scale (SUS) questions for the Active and Independent Management System app (n=38).

Statements	Participants agreeing, n (%)	Participants disagreeing, n (%)	SUS scores, mean (SD)
Positive statements			
I think that I would like to use this system frequently	34 (89)	1 (3)	4.16 (0.68)
I thought the system was easy to use	31 (82)	1 (3)	4.13 (0.62)
I found the various functions in this system were well integrated	34 (89)	1 (3)	4.16 (0.68)
I would imagine that most people would learn to use this system very quickly	32 (84)	2 (5)	3.92 (0.67)
I felt very confident using this system	33 (87)	2 (5)	3.97 (0.68)
Negative statements			
I found the system unnecessarily complex	1 (3)	32 (84)	2.00 (0.66)
I think that I would need the support of a technical person to be able to use this system	1 (3)	32 (84)	1.92 (0.71)
I thought there was too much inconsistency in this system	1 (3)	32 (84)	2.11 (0.56)
I found the system very cumbersome to use	1 (3)	28 (74)	2.24 (0.59)
I needed to learn a lot of things before I could get going with this system	2 (5)	26 (68)	2.32 (0.66)

AIMS App: UEQ Results

Invitations were sent to 12 (27%) of the 44 participants. Of these 12 patients, 10 (83%) completed the UEQ questionnaire. The means of the normalized scores (range −3 to +3) for the AIMS app were as follows: attractiveness=2.733 (SD 0.070), perspicuity=2.900 (SD 0.060), efficiency=2.800 (SD 0.090), dependability=2.425 (SD 0.060), stimulation=2.200 (SD 0.010), and novelty=1.450 (0.260). Figure 5 shows the bar chart of the results for the AIMS app against the benchmarks, with all dimensions classed as *excellent* (with the exception of *novelty*, which was classed as *good*), providing slightly better results than the SUS questionnaire. Table 4 gives the mean (SD) and variance of the normalized values for the items in the UEQ questionnaire for the AIMS app. In this case, all values are very encouraging. Figure 6 shows the bar chart for the data grouped into the 6 UEQ dimensions.

Figure 5. Bar chart of the Active and Independent Management System app User Experience Questionnaire results against the benchmarks.

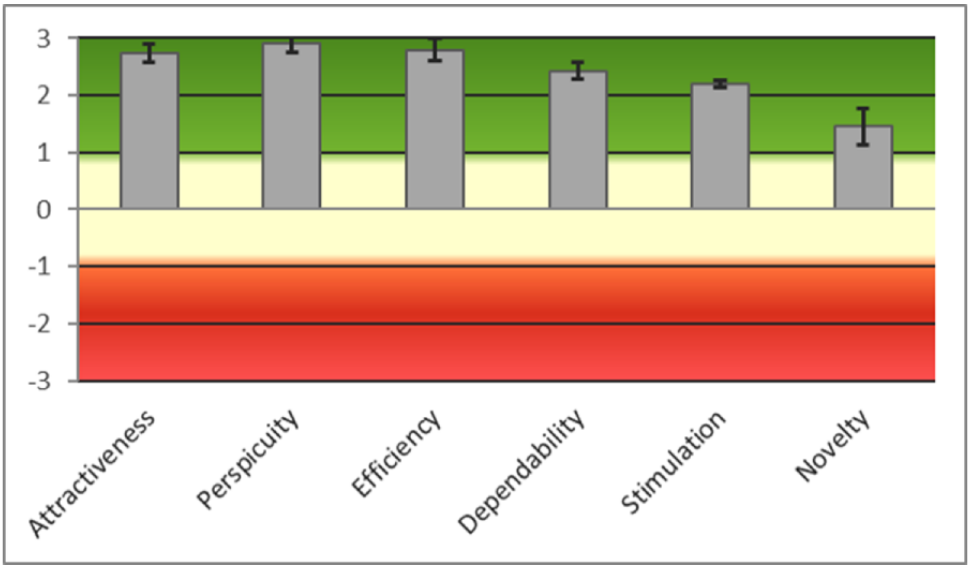


XSL•FO
RenderX

Table 4. Mean (SD) and variance of the normalized values for the items in the User Experience Questionnaire (UEQ) for the Active and Independent Management System app (n=10).

Scale	Left anchor of the scale	Right anchor of the scale	UEQ scores, mean (SD)	Variance
Attractiveness	Annoying	Enjoyable	1.9 (0.3)	0.1
Perspicuity	Not understandable	Understandable	2.6 (0.7)	0.5
Novelty	Creative	Dull	1.9 (0.3)	0.1
Perspicuity	Easy to learn	Difficult to learn	2.9 (0.3)	0.1
Stimulation	Valuable	Inferior	2.2 (0.6)	0.4
Stimulation	Boring	Exciting	1.5 (0.5)	0.3
Stimulation	Not interesting	Interesting	2.1 (0.6)	0.3
Dependability	Unpredictable	Predictable	1.8 (0.4)	0.2
Efficiency	Fast	Slow	2.8 (0.4)	0.2
Novelty	Inventive	Conventional	1.9 (0.3)	0.1
Dependability	Obstructive	Supportive	2.2 (0.6)	0.4
Attractiveness	Good	Bad	2.8 (0.4)	0.2
Perspicuity	Complicated	Easy	2.7 (0.5)	0.2
Attractiveness	Unlikable	Pleasing	2.7 (0.5)	0.2
Novelty	Usual	Leading edge	1.2 (0.4)	0.2
Attractiveness	Unpleasant	Pleasant	2.9 (0.3)	0.1
Dependability	Secure	Not secure	2.5 (0.5)	0.3
Stimulation	Motivating	Demotivating	2.0 (0.5)	0.2
Dependability	Meets expectations	Does not meet expectations	2.7 (0.5)	0.2
Efficiency	Inefficient	Efficient	2.8 (0.4)	0.2
Perspicuity	Clear	Confusing	2.9 (0.3)	0.1
Efficiency	Impractical	Practical	2.7 (0.5)	0.2
Efficiency	Organized	Cluttered	2.8 (0.4)	0.2
Attractiveness	Attractive	Unattractive	2.9 (0.3)	0.1
Attractiveness	Friendly	Unfriendly	2.9 (0.3)	0.1
Novelty	Conservative	Innovative	1.5 (0.5)	0.3

Figure 6. Bar chart of the Active and Independent Management System app data grouped into the 6 User Experience Questionnaire dimensions.



XSL•FO
RenderX

AIMS App: Qualitative Feedback

To gain further insight into how users perceived the AIMS app, 3 additional questions were asked (refer to the following subsections).

Q1. What Do You Think Are the Advantages of This App?

Most of the participants (33/44, 75%) answered this question. Of the 33 participants, 26 (79%) thought that the exercise videos provided after an operation were very useful, 30 (91%) considered clinicians having immediate access to patient progress an advantage, and 24 (73%) considered ease of use an advantage. Example comments were as follows:

> *Having exercise videos that I can use both in the hospital and at home is a great help. While there is help on hand in the hospital if needed, being able to view the videos while at home is great.* [Patient A]
>
> *Loved the being able to access the videos on the tablet, was very helpful and the app was so easy to use.* [Patient B]

Q2. What Do You Think Are the Disadvantages of This App?

Only 8 (18%) of the 44 participants answered this question, and very few disadvantages were listed: 1 (13%) participant thought that the app could include some embedded videos for generic stretching exercises; 1 (13%) thought that the app might be too simple, and more functionality was required; and 6 (75%) thought that a self-help section would be beneficial. An example comment was as follows:

> *While the hospital provide[s] a number of leaflets on what to expect after the knee replacement, it would be handier of [sic] these were part of the app.* [Patient C]

Q3. Would You Change Anything?

Of the 44 participants, 17 (39%) answered this question. Of these 17 participants, 6 (35%) suggested more self-help, and 5 (29%) suggested having the ability to keep a daily or weekly diary of symptoms or pain. An example comment was as follows:

> *Would it be possible to have a section in the app to record how I am getting on with the videos and make notes on any symptoms I'm getting after the operation, particularly once I'm home?* [Patient D]

Discussion

Principal Findings

The main goal of this mixed methods study was to examine the usability of the AIMS platform from the perspectives of both patients and clinicians. Two well-known validated instruments were used to measure usability: the SUS and the UEQ. In all, 26 physiotherapists and health care professionals evaluated the AIMS clinical portal; and 44 patients in hospital for TKR, THR, or DHS evaluated the AIMS app. In terms of the RQs, the study has shown that both the AIMS clinical portal (RQ1) and the AIMS app (RQ2) have *good* to *excellent* usability scores, and this platform provides a solid foundation for the next phase of research, which will involve evaluating its effectiveness in improving patient outcomes after TKR, THR, or DHS. In addition, useful qualitative information was obtained from participants through a set of open-ended questions.

On the basis of the literature reviewed in the Re-Enablement DHIs subsection, it seems that smartphones and the web are the 2 main platforms used to provide re-enablement DHIs after TKA or THA. The platforms have been identified to be used by patients who will receive the instructions in the form of video, text, and interactive game as well as by clinicians who can create custom treatment plans for patients. The AIMS platform provides similar functionality to the systems found in the literature, with a web-based clinical portal and a mobile app for patients. The AIMS platform mainly presents content in video and text, which is similar to the majority of the systems discussed in the Re-Enablement DHIs subsection. Text and video are considered to be effective in presenting rehabilitation content to patients because they allow a wider level of proficiency in information and communications technology. Compared to static images, we considered videos to be more engaging, although further research should be conducted to investigate this. Some studies, such as those by Hussain et al [23], van Dijk-Huisman et al [25], and Bell et al [27], used sensors from wearable devices and mobile phones, while Bäcker et al [28] developed their own custom sensor. Personalization features that allow the system to customize activities for patients were only evident in the studies by Hussain et al [23], van Dijk-Huisman et al [25], and Bell et al [27]. Currently, the AIMS platform does not use sensors or have any personalization features, but these will be considered for the next phase of the research. A summary comparing the literature reviewed with our study can be found in Multimedia Appendix 1 [23-30].

In terms of limitations, to overcome major privacy concerns, a condition of the ethics approval for the study was that all data had to be anonymized; therefore, neither could we perform a demographic analysis nor conduct follow-up monitoring of the progress of a more informed patient after they left the hospital. As this study's main focus was on usability, the recruitment of participating patients was carried out during rehabilitation sessions. This method did not allow us to conduct a randomized study, and there were no control and experimental groups. Furthermore, due to ethics approval restrictions, we were not able to directly observe the experiment and had to use questionnaires and interviews conducted by the clinicians during the rehabilitation sessions. The experiment did not use any additional sensing technologies to monitor user progress and relied on the patient's input and feedback. Future studies will aim to overcome these limitations.

Since this study was carried out, the platform has been improved to include additional support videos for patients, ideas for which emerged from the qualitative feedback, and a second usability study is underway to ensure that results are consistent with this initial study (Multimedia Appendix 1).

XSL•FO
RenderX

Conclusions

This study aimed to assess the usability of a re-enablement platform called AIMS, designed to address the *de-enablement* often experienced by hospitalized older adults most at risk. Usability was measured using 2 common validated instruments: the 10-item SUS and, for more detailed analysis, the 26-item UEQ. The AIMS clinical portal was evaluated by 26 physiotherapists and health care professionals; and 44 patients undergoing TKR, THR, or DHS assessed the AIMS app. Overall, both the AIMS clinical portal and the AIMS app received *good* to *excellent* usability scores, providing a solid foundation for future research on their effectiveness in improving patient outcomes after joint replacements. Optimal rehabilitation programs for orthopedic joint replacement patients can lead to a quicker return to normal function, faster hospital discharge, and higher patient satisfaction.

Acknowledgments

The authors acknowledge the resources allocated to this study by Hairmyres University Hospital, Lanarkshire.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Comparison with the literature reviewed.
[XLSX File (Microsoft Excel File), 6 KB - humanfactors_v11i1e50430_app1.xlsx]

References

1. Mackey S, Gilam G, Darnall B, Goldin P, Kong JT, Law C, et al. Mindfulness-based stress reduction, cognitive behavioral therapy, and acupuncture in chronic low back pain: protocol for two linked randomized controlled trials. JMIR Res Protoc 2022 Sep 27;11(9):e37823 [FREE Full text] [doi: 10.2196/37823] [Medline: 36166279]
2. The state of musculoskeletal health. Versus Arthritis. URL: https://www.versusarthritis.org/about-arthritis/data-and-statistics/the-state-of-musculoskeletal-health/ [accessed 2022-11-03]
3. Hewitt S, Sephton R, Yeowell G. The effectiveness of digital health interventions in the management of musculoskeletal conditions: systematic literature review. J Med Internet Res 2020 Jun 05;22(6):e15617 [FREE Full text] [doi: 10.2196/15617] [Medline: 32501277]
4. Nagel J, Wegener F, Grim C, Hoppe MW. Effects of digital physical health exercises on musculoskeletal diseases: systematic review with best-evidence synthesis. JMIR Mhealth Uhealth 2024 Jan 23;12:e50616 [FREE Full text] [doi: 10.2196/50616] [Medline: 38261356]
5. Briggs AM, Woolf AD, Dreinhöfer K, Homb N, Hoy DG, Kopansky-Giles D, et al. Reducing the global burden of musculoskeletal conditions. Bull World Health Organ 2018 Apr 12;96(5):366-368. [doi: 10.2471/blt.17.204891]
6. Duffield SJ, Ellis BM, Goodson N, Walker-Bone K, Conaghan PG, Margham T, et al. The contribution of musculoskeletal disorders in multimorbidity: implications for practice and policy. Best Pract Res Clin Rheumatol 2017 Apr;31(2):129-144 [FREE Full text] [doi: 10.1016/j.berh.2017.09.004] [Medline: 29224692]
7. Ross EL, Jamison RN, Nicholls L, Perry BM, Nolen KD. Clinical integration of a smartphone app for patients with chronic pain: retrospective analysis of predictors of benefits and patient engagement between clinic visits. J Med Internet Res 2020 Apr 16;22(4):e16939 [FREE Full text] [doi: 10.2196/16939] [Medline: 32297871]
8. Williams A, Kamper SJ, Wiggers JH, O'Brien KM, Lee H, Wolfenden L, et al. Musculoskeletal conditions may increase the risk of chronic disease: a systematic review and meta-analysis of cohort studies. BMC Med 2018 Sep 25;16(1):167 [FREE Full text] [doi: 10.1186/s12916-018-1151-2] [Medline: 30249247]
9. Keavy R. The prevalence of musculoskeletal presentations in general practice: an epidemiological study. Br J Gen Pract 2020 Jun 18;70(suppl 1):bjgp20X711497. [doi: 10.3399/bjgp20x711497]
10. Health matters: health and work. United Kingdom Government. URL: https://www.gov.uk/government/publications/health-matters-health-and-work/health-matters-health-and-work [accessed 2022-11-02]
11. mHealth: use of appropriate digital technologies for public health: report by the Director-General. World Health Organization. 2017. URL: https://iris.who.int/handle/10665/274134 [accessed 2024-04-16]
12. Cunningham SG, Wake DJ, Waller A, Morris AD, Walker J. My diabetes my way: an electronic personal health record for diabetes. Diabetes Vasc Dis 2013 Jun 26;13(3):143-149. [doi: 10.1177/1474651413493336]
13. Janssen H, Cobbe J, Norval C, Singh J. Decentralized data processing: personal data stores and the GDPR. Int Data Privacy Law 2020 Nov;10(4):356-384 [FREE Full text] [doi: 10.1093/idpl/ipaa016]
14. Wherton J, Greenhalgh T, Shaw SE. Expanding video consultation services at pace and scale in Scotland during the COVID-19 pandemic: national mixed methods case study. J Med Internet Res 2021 Oct 07;23(10):e31374 [FREE Full text] [doi: 10.2196/31374] [Medline: 34516389]

XSL•FO
RenderX

15. Murray E, Hekler EB, Andersson G, Collins LM, Doherty A, Hollis C, et al. Evaluating digital health interventions: key questions and approaches. Am J Prev Med 2016 Nov;51(5):843-851 [FREE Full text] [doi: 10.1016/j.amepre.2016.06.008] [Medline: 27745684]
16. Garell C, Svedberg P, Nygren JM. A legal framework to support development and assessment of digital health services. JMIR Med Inform 2016 May 25;4(2):e17 [FREE Full text] [doi: 10.2196/medinform.5401] [Medline: 27226391]
17. Classification of digital health interventions v1.0: a shared language to describe the uses of digital technology for health. World Health Organization. URL: https://apps.who.int/iris/bitstream/handle/10665/260480/WHO-RHR-18.06-eng.pdf [accessed 2022-12-03]
18. Saleem M, Kühne L, De Santis KK, Christianson L, Brand T, Busse H. Understanding engagement strategies in digital interventions for mental health promotion: scoping review. JMIR Ment Health 2021 Dec 20;8(12):e30000 [FREE Full text] [doi: 10.2196/30000] [Medline: 34931995]
19. Hardy A, Courgeon M, Pellei K, Desmeules F, Loubert C, Vendittoli PA. Improved clinical outcomes of outpatient enhanced recovery hip and knee replacements in comparison to standard inpatient procedures: a study of patients who experienced both. Orthop Traumatol Surg Res 2022 Oct;108(6):103236 [FREE Full text] [doi: 10.1016/j.otsr.2022.103236] [Medline: 35150928]
20. Zhong M, Liu D, Tang H, Zheng Y, Bai Y, Liang Q, et al. Impacts of the perioperative fast track surgery concept on the physical and psychological rehabilitation of total hip arthroplasty: a prospective cohort study of 348 patients. Medicine (Baltimore) 2021 Aug 13;100(32):e26869 [FREE Full text] [doi: 10.1097/MD.0000000000026869] [Medline: 34397901]
21. Galbraith AS, McGloughlin E, Cashman J. Enhanced recovery protocols in total joint arthroplasty: a review of the literature and their implementation. Ir J Med Sci 2018 Feb 16;187(1):97-109. [doi: 10.1007/s11845-017-1641-9] [Medline: 28623570]
22. Swinnerton E, Price A. Recognising, reducing and preventing deconditioning in hospitalised older people. Nurs Older People 2023 Apr 03;35(2):34-41. [doi: 10.7748/nop.2023.e1396] [Medline: 36751897]
23. Hussain MS, Li J, Brindal E, van Kasteren Y, Varnfield M, Reeson A, et al. Supporting the delivery of total knee replacements care for both patients and their clinicians with a mobile app and web-based tool: randomized controlled trial protocol. JMIR Res Protoc 2017 Mar 01;6(3):e32 [FREE Full text] [doi: 10.2196/resprot.6498] [Medline: 28249832]
24. Timmers T, Janssen L, van der Weegen WJ, Das D, Marijnissen W, Hannink G, et al. The effect of an app for day-to-day postoperative care education on patients with total knee replacement: randomized controlled trial. JMIR Mhealth Uhealth 2019 Oct 21;7(10):e15323 [FREE Full text] [doi: 10.2196/15323] [Medline: 31638594]
25. van Dijk-Huisman HC, Weemaes AT, Boymans TA, Lenssen AF, de Bie RA. Smartphone app with an accelerometer enhances patients' physical activity following elective orthopedic surgery: a pilot study. Sensors (Basel) 2020 Aug 02;20(15):4317 [FREE Full text] [doi: 10.3390/s20154317] [Medline: 32748876]
26. Wijnen A, Hoogland J, Munsterman T, Gerritsma CL, Dijkstra B, Zijlstra WP, et al. Effectiveness of a home-based rehabilitation program after total hip arthroplasty driven by a tablet app and remote coaching: nonrandomized controlled trial combining a single-arm intervention cohort with historical controls. JMIR Rehabil Assist Technol 2020 Apr 27;7(1):e14139 [FREE Full text] [doi: 10.2196/14139] [Medline: 32338621]
27. Bell KM, Onyeukwu C, Smith CN, Oh A, Devito Dabbs A, Piva SR, et al. A portable system for remote rehabilitation following a total knee replacement: a pilot randomized controlled clinical study. Sensors (Basel) 2020 Oct 27;20(21):6118 [FREE Full text] [doi: 10.3390/s20216118] [Medline: 33121204]
28. Bäcker HC, Wu CH, Schulz MR, Weber-Spickschen TS, Perka C, Hardt S. App-based rehabilitation program after total knee arthroplasty: a randomized controlled trial. Arch Orthop Trauma Surg 2021 Sep 06;141(9):1575-1582 [FREE Full text] [doi: 10.1007/s00402-021-03789-0] [Medline: 33547927]
29. Colomina J, Drudis R, Torra M, Pallisó F, Massip M, Vargiu E, et al. Implementing mHealth-enabled integrated care for complex chronic patients with osteoarthritis undergoing primary hip or knee arthroplasty: prospective, two-arm, parallel trial. J Med Internet Res 2021 Sep 02;23(9):e28320 [FREE Full text] [doi: 10.2196/28320] [Medline: 34473068]
30. Rian T, Sand K, Skogvoll E, Klepstad P, Wik TS. A web-based communication tool for postoperative follow-up and pain assessment at home after primary knee arthroplasty: feasibility and usability study. JMIR Form Res 2022 Apr 28;6(4):e34543 [FREE Full text] [doi: 10.2196/34543] [Medline: 35482392]
31. Bäcker HC, Wu CH, Pförringer D, Petersen W, Stöckle U, Braun KF. A review of functional outcomes after the app-based rehabilitation of patients with TKA and THA. J Pers Med 2022 Aug 21;12(8):1342 [FREE Full text] [doi: 10.3390/jpm12081342] [Medline: 36013291]
32. Constantinescu D, Pavlis W, Rizzo M, Vanden Berge D, Barnhill S, Hernandez VH. The role of commercially available smartphone apps and wearable devices in monitoring patients after total knee arthroplasty: a systematic review. EFORT Open Rev 2022 Jul 05;7(7):481-490 [FREE Full text] [doi: 10.1530/EOR-21-0115] [Medline: 35900191]
33. Brooke J. SUS: a 'quick and dirty' usability scale. In: Usability Evaluation In Industry. Boca Raton, FL: CRC Press; 1996.
34. Laugwitz B, Held T, Schrepp M. Construction and evaluation of a user experience questionnaire. In: Proceedings of the 4th Symposium of the Workgroup Human-Computer Interaction and Usability Engineering of the Austrian Computer Society. 2008 Presented at: USAB 2008; November 20-21, 2008; Graz, Austria. [doi: 10.1007/978-3-540-89350-9_6]
35. WHO evaluation practice handbook. World Health Organization. 2013. URL: http://apps.who.int/iris/bitstream/10665/96311/1/9789241548687_eng.pdf [accessed 2023-01-04]

XSL•FO
RenderX

36. Seffah A, Donyaee M, Kline RB, Padda HK. Usability measurement and metrics: a consolidated model. Softw Qual J 2006 Jun;14(2):159-178. [doi: 10.1007/s11219-006-7600-8]
37. BS EN ISO 9241-11:1998 ergonomic requirements for office work with visual display terminals (VDTs) - Guidance on usability. British Standards Institution. 1998. URL: https://knowledge.bsigroup.com/products/ergonomic-requirements-for-office-work-with-visual-display-terminals-vdts-guidance-on-usability?version=standard [accessed 2024-04-16]
38. ISO/IEC 9126 in software engineering. GeeksforGeeks. URL: https://www.geeksforgeeks.org/iso-iec-9126-in-software-engineering/ [accessed 2024-04-16]
39. Hornbæk K. Current practice in measuring usability: challenges to usability studies and research. Int J Hum Comput Stud 2006 Feb;64(2):79-102. [doi: 10.1016/j.ijhcs.2005.06.002]
40. Johnson SG, Potrebny T, Larun L, Ciliska D, Olsen NR. Usability methods and attributes reported in usability studies of mobile apps for health care education: scoping review. JMIR Med Educ 2022 Jun 29;8(2):e38259 [FREE Full text] [doi: 10.2196/38259] [Medline: 35767323]
41. Bangor A, Kortum PT, Miller JT. An empirical evaluation of the system usability scale. Int J Hum Comput Interact 2008 Jul 30;24(6):574-594. [doi: 10.1080/10447310802205776]
42. Schrepp M, Hinderks A, Thomaschewski J. Applying the user experience questionnaire (UEQ) in different evaluation scenarios. In: Proceedings of the Design, User Experience, and Usability: Theories, Methods, and Tools for Designing the User Experience. 2014 Presented at: DUXU 2014; June 22-27, 2014; Heraklion, Greece. [doi: 10.1007/978-3-319-07668-3_37]

Abbreviations

AIMS: Active and Independent Management System
DHI: digital health intervention
DHS: dynamic hip screw
IC: integrated care
mHealth: mobile health
POD1: postoperative day 1
RQ: research question
SUS: System Usability Scale
THA: total hip arthroplasty
THR: total hip replacement
TKA: total knee arthroplasty
TKR: total knee replacement
UEQ: User Experience Questionnaire

Edited by A Parush; submitted 30.06.23; peer-reviewed by C Baxter, F Hussain; comments to author 14.01.24; revised version received 08.03.24; accepted 18.03.24; published 14.05.24.

Please cite as:
Papadopoulos P, Soflano M, Connolly T
A Digital Health Intervention Platform (Active and Independent Management System) to Enhance the Rehabilitation Experience for Orthopedic Joint Replacement Patients: Usability Evaluation Study
JMIR Hum Factors 2024;11:e50430
URL: https://humanfactors.jmir.org/2024/1/e50430
doi:10.2196/50430
PMID:38743479

©Petros Papadopoulos, Mario Soflano, Thomas Connolly. Originally published in JMIR Human Factors (https://humanfactors.jmir.org), 14.05.2024. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on https://humanfactors.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Evaluation of the Parkinson's Remote Interactive Monitoring System in a Clinical Setting: Usability Study

Bronwyn Bridges[1], BSc; Jake Taylor[2], BSc; John Thomas Weber[1], BS, MS, PhD

[1]School of Pharmacy, Memorial University, St. John's, NL, Canada

[2]School of Exercise Science, Physical & Health Education, University of Victoria, Victoria, BC, Canada

Corresponding Author:
John Thomas Weber, BS, MS, PhD
School of Pharmacy
Memorial University
Health Sciences Centre
300 Prince Philip Drive
St. John's, NL, A1B 3V6
Canada
Phone: 1 709 864 7261
Fax: 1 709 864 6941
Email: jweber@mun.ca

Abstract

Background: The fastest-growing neurological disorder is Parkinson disease (PD), a progressive neurodegenerative disease that affects 10 million people worldwide. PD is typically treated with levodopa, an oral pill taken to increase dopamine levels, and other dopaminergic agonists. As the disease advances, the efficacy of the drug diminishes, necessitating adjustments in treatment dosage according to the patient's symptoms and disease progression. Therefore, remote monitoring systems that can provide more detailed and accurate information on a patient's condition regularly are a valuable tool for clinicians and patients to manage their medication. The Parkinson's Remote Interactive Monitoring System (PRIMS), developed by PragmaClin Research Inc, was designed on the premise that it will be an easy-to-use digital system that can accurately capture motor and nonmotor symptoms of PD remotely.

Objective: We performed a usability evaluation in a simulated clinical environment to assess the ease of use of the PRIMS and determine whether the product offers suitable functionality for users in a clinical setting.

Methods: Participants were recruited from a user sign-up web-based database owned by PragmaClin Research Inc. A total of 11 participants were included in the study based on the following criteria: (1) being diagnosed with PD and (2) not being diagnosed with dementia or any other comorbidities that would make it difficult to complete the PRIMS assessment safely and independently. Patient users completed a questionnaire that is based on the Movement Disorder Society–sponsored revision of the Unified Parkinson's Disease Rating Scale. Interviews and field notes were analyzed for underlying themes and topics.

Results: In total, 11 people with PD participated in the study (female individuals: n=5, 45%; male individuals: n=6, 55%; age: mean 66.7, SD 7.77 years). Thematic analysis of the observer's notes revealed 6 central usability issues associated with the PRIMS. These were the following: (1) the automated voice prompts are confusing, (2) the small camera is problematic, (3) the motor test exhibits excessive sensitivity to the participant's orientation and position in relation to the cameras, (4) the system poses mobility challenges, (5) navigating the system is difficult, and (6) the motor test exhibits inconsistencies and technical issues. Thematic analysis of qualitative interview responses revealed four central themes associated with participants' perspectives and opinions on the PRIMS, which were (1) admiration of purpose, (2) excessive system sensitivity, (3) video instructions preferred, and (4) written instructions disliked. The average system usability score was calculated to be 69.2 (SD 4.92), which failed to meet the acceptable system usability score of 70.

Conclusions: Although multiple areas of improvement were identified, most of the participants showed an affinity for the overarching objective of the PRIMS. This feedback is being used to upgrade the current PRIMS so that it aligns more with patients' needs.

(JMIR Hum Factors 2024;11:e54145) doi:10.2196/54145

XSL•FO
RenderX

KEYWORDS

Parkinson disease; usability; remote monitoring; motor examination; movement disorders; thematic analysis; System Usability Scale; mobile phone

Introduction

Background

Overview

The fastest-growing neurological disorder is Parkinson disease (PD) [1]. PD is a progressive neurodegenerative disease that affects 10 million people worldwide [2]. The incidence and prevalence of PD is rising sharply in countries with aging populations, and in the last 2 decades, the burden of PD has more than doubled, with estimations predicting 1,238,000 cases in North America by 2023 [3,4]. The disease affects the basal nuclei in the central nervous system causing the progressive deterioration of dopaminergic neurons. The loss of these neurons causes motor and nonmotor dysfunctions [5]. Motor system deficits result in symptoms such as tremor, rigidity, bradykinesia, and postural instability [6]. Other symptoms include cognitive problems, gastrointestinal upset, and urinary control issues [7]. Due to these mal effects, PD is linked to morbidity, high economic burden, and decreased quality of life for patients and caregivers. The annual estimated direct and indirect costs of the condition in the United States alone are close to US $52 billion [8]. Neurologists are struggling to manage the increasing prevalence of PD, leading to clinician burnout [9] and lengthy appointment wait times for patients [10]. However, studies show that the management of these symptoms in the early stages of the disease can achieve positive results. In contrast, the consequences of late or faulty diagnoses negatively impact patients and the health care system [11-13].

Medication Management

PD is typically treated with levodopa, an oral pill taken to increase dopamine levels, and other dopaminergic agonists. However, as the disease progresses, the effects of the drugs wane. This requires medication dosage adjustments to properly manage symptoms throughout the day [14]. This can be a difficult task for physicians as symptoms are constantly fluctuating and may appear and disappear throughout the day with a hard-to-establish pattern. Some physicians ask their patients to keep diaries where they note the time of day and a description of their symptoms. However, adherence to this method is typically poor and does not provide meaningful information [15]. Therefore, remote monitoring systems that provide more detailed and accurate information on a patient's condition regularly are a valuable tool for clinicians and patients to manage their medication.

Evaluation of PD

The evaluation of PD is commonly performed using clinical rating scales that are essential to the quantification of neurological disorders [16]. These rating scales enable clinicians and researchers to evaluate PD symptoms, progression, treatment efficacy, and disease severity [16,17]. One of the most widely used clinical scales for PD assessment is the Movement Disorder Society–sponsored revision of the Unified Parkinson's Disease Rating Scale (MDS-UPDRS) [18].

The MDS-UPDRS is a revised form of the original Unified Parkinson's Disease Rating Scale [17] and incorporates both motor and nonmotor aspects into the assessment. It consists of 65 elements and, on average, requires approximately 30 minutes of administration time. There are four parts to the questionnaire: (1) nonmotor experiences of daily living (13 elements), (2) motor experiences of daily living (13 elements), (3) motor examination (33 elements), and (4) motor complications (6 elements) [18]. Elements are scored from 0 to 4, where 0=normal, 1=slight, 2=mild, 3=moderate, and 4=severe. There are some elements that patients could possibly administer themselves as they are multiple-choice questions asking about personal symptom experience, whereas others are rated by an examiner (typically a neurologist or other clinician) based on observation and physical examination.

While the MDS-UPDRS represents the international gold standard in PD rating scales and has undergone strict validation through clinical studies, it still remains a clinician-based scale; this means that a clinician assigns a score based on their own personal qualitative observations of a patient. Therefore, the assessments are often subjective and biased to the examiner's skill and knowledge. The assessments will also vary from one examiner to the other in this way [19-21]. Studies have shown that there is variability between assessments conducted by nurses and neurologists [22,23]. In these situations, it is difficult to compare and interpret the scores, as they may differ based on a patient's condition or simply due to the clinician performing the assessment. The MDS-UPDRS is also time consuming for clinicians. It requires approximately 30 minutes of an examiner's time, which makes it impractical for routine practice [18]. Examiners must also be highly trained to improve the validity of the scores. Many of the elements in the MDS-UPDRS must be completed by a patient themselves, which adds to the time burden of the questionnaire when performed in a clinician's office. The typical assessment performed in a clinical setting rarely assesses a patient's day-to-day symptoms, which usually vary over time, and only captures a snapshot of an individual's condition at the moment of their appointment [24]. Patients also typically have long wait times in between their appointments, which makes it difficult to remember their symptoms since their last visit [10]. This way, medical decisions are now influenced by recall bias and patient attitudes instead of by reliable patient data. In addition, it is an inconvenience for patients to travel to clinics due to transportation, long commutes, and their conditions, especially if they are in the advanced stages of PD. Therefore, there is a need for objective, accurate, and reliable assessment tools that can help increase the chance of effective treatment. These could aid patients with their disease management, thus cutting down on health care costs [25].

XSL•FO
RenderX

Digital Health Technologies and the Parkinson's Remote Interactive Monitoring System

An emerging solution to some access to health care issues are video-based visits. These bring care directly into a patient's home, which improves access in a patient-centered manner and minimizes the burden on people with PD and their caregivers [26]. In addition, due to the largely visual nature of a PD examination, it tends to work well in a video-based visit. Studies have shown that web-based appointments with neurologists are feasible and valuable [27,28]. It has also been shown that a modified version of the MDS-UPDRS motor examination (excluding the test of rigidity and postural stability) can be successfully administered remotely [29]. However, as virtual visits still require a clinician's time, they still only provide a brief snapshot of a patient's condition. There need to be other methods of assessing PD without occupying already overburdened clinicians.

Digital health technologies that alleviate the need for medical professionals to assess disease progression have been on the rise. These technologies offer possibilities for self-assessment and improved health care [30]. Some of the technologies developed for PD include wearable sensors and mobile apps. These devices have been used extensively to monitor motor symptoms and complications of people with PD in their home environments [31]. These wearable sensors and mobile apps can accurately track the progression of PD [32-34] and other neurological conditions [35]. Examples of these devices on the market are the Global Kinetics Corporation's Personal KinetiGraph Watch [36,37] and Rune Labs' StrivePD mobile app [38]. APDM Wearable Technologies has also developed multiple sensors that can accurately monitor tremor and dyskinesia symptoms of PD that have been used in many clinical studies [32,34,39,40]. Other wearables that collect contextual data include DynaPort MiniMod Hybrid (a sensor worn on the lower back), Shimmer (records gait), SENSE-PARK (records walking, hypokinesia, dyskinesia, and sleep), activPAL, and StepWatch (gait and basic movement parameters) [20,41-45]. The problem is that these technologies only generate a small amount of patient data (mainly tremors and other motor symptoms, moods, and sleep characteristics). Therefore, although these devices provide an objective means of tracking PD characteristics, they do not provide a complete assessment of the condition. Wearable sensors also have inherent risks [46] and do not follow the gold standard clinical scales such as the MDS-UPDRS. These risks encompass potential interference with the daily activities of patients with PD, impacting their natural movements and behaviors. In addition, behavioral modifications stemming from the feedback provided by sensors can yield both positive and negative outcomes. On the positive side, such modifications may encourage beneficial lifestyle changes and provide meaningful data. However, as a downside, they may also contribute to increased anxiety and foster a dependency on the wearable device [46].

To address the need for reliable tools to objectively assess PD symptoms that do not require a clinician's involvement, the Parkinson's Remote Interactive Monitoring System (PRIMS) was developed. The PRIMS is a digitized version of the MDS-UPDRS in the form of a desktop application that people with PD can complete themselves. This way, the PRIMS provides a complete picture of PD assessment via its capacity to comprehensively measure both motor and nonmotor symptoms without the need for wearable sensors and its potential to serve as a valuable tool in a clinical or home setting. If validated through further investigation, the PRIMS has the potential of delivering a standard in PD assessment. The PRIMS also has the potential to be valuable in a home setting, offering a user-operated system capable of capturing a significant portion of the MDS-UPDRS (considered the gold standard). The system provides patients with a means of tracking their condition remotely and offers clinicians reliable data for better medication management. This comprehensive approach enhances understanding and facilitates more effective monitoring of the progression and individual symptoms of a patient.

Usability Testing

The development of any system that is used by patients and clinicians for the management of biomedical data should always involve usability evaluations, which aim to understand whether such a product is easy to use and has the appropriate functionality for the users. *Usability* is a term used to define how easily people can use a tool or object to accomplish a specific task [47,48]. In this way, when developing interfaces, it is imperative that they can be learned quickly and are easy to navigate. The system's layout should avoid and manage operational errors efficiently and provide users with appropriate feedback [47]. Usability must also address user satisfaction and provide solutions to the problem that the system was designed to solve [49]. A common method of assessing usability is the System Usability Scale (SUS). The SUS has been used in multiple studies, such as the evaluation of a mobile app for people with PD [50]. Structured interviews are common practice for these types of studies [51]. Field notes can also be a valuable tool for qualitative researchers to collect and analyze [52]. Observational notes can capture information such as the nonverbal reactions of users while they interact with the system. This study used multiple methods to assess the usability of the PRIMS.

Study Objectives

This study aimed to assess the functionality, usability, and user experience aspects of the most recent version of the PRIMS in a clinical setting from the perspectives of people with PD. Use issues identified in this study will guide designers in creating a more effective commercial product. Using multiple methods, including interviews and field notes along with SUS surveys, we evaluated the user experience of the PRIMS.

Methods

Participants

Participants were recruited from a user sign-up web-based database owned by PragmaClin Research Inc. The study was also advertised by the Parkinson Society Newfoundland and Labrador on their weekly newsletter. Interested participants who contacted us were given a questionnaire that determined their eligibility for the study. The inclusion criteria for study participation were the following: (1) being diagnosed with PD

XSL•FO
RenderX

and (2) not being diagnosed with dementia or any other comorbidities that would make it difficult to complete the PRIMS assessment safely and independently. Participants were recruited on a first come, first served basis. Informed consent was obtained from all participants via a web-based consent form emailed to them before study completion. Paper copies were also available to participants at the time of their scheduled session.

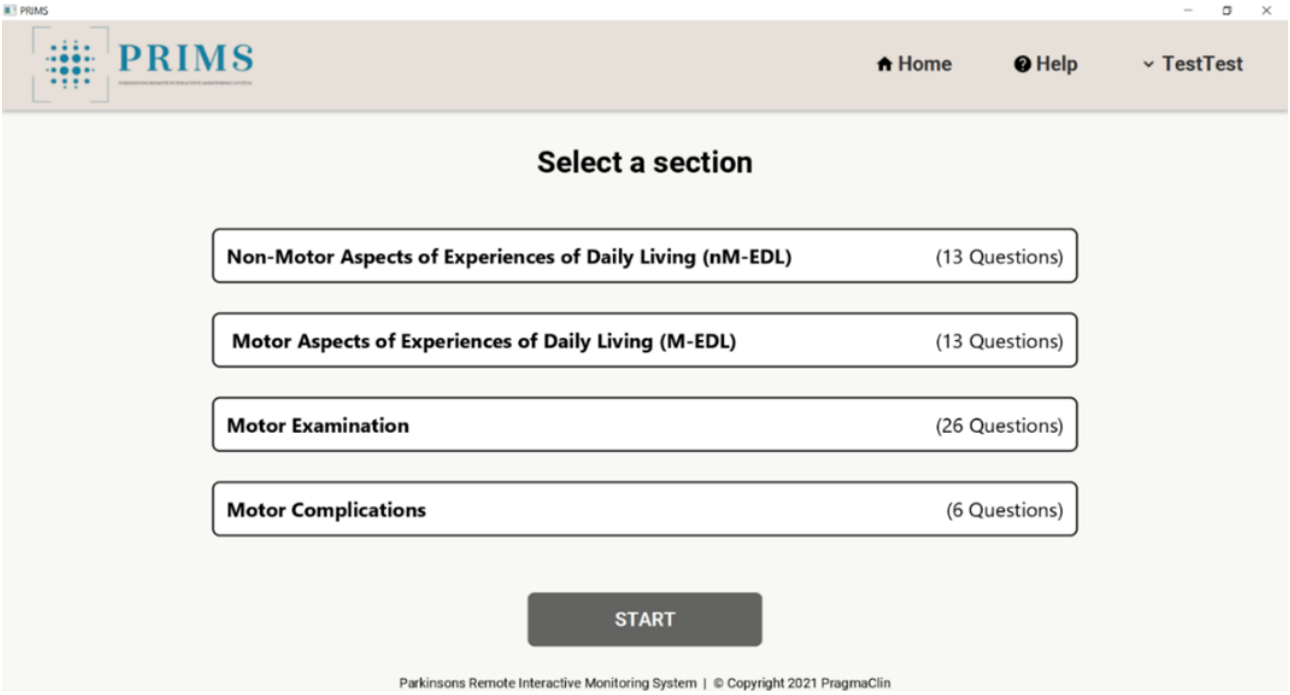
Ethical Considerations

This study received ethics approval from the National Research Council of Canada Institutional Review Board (protocol 2021-137). Informed consent was obtained, and the possible consequences of the study were explained. All data were deidentified. No compensation was provided to participants.

Description of the PRIMS

The PRIMS was developed by PragmaClin Research Inc and was designed on the premise that it will be an easy-to-use digital system that can accurately quantify motor and nonmotor symptoms of PD remotely. The PRIMS has the capability to interact with patients in real time, delivering results promptly through a dedicated patient dashboard. Patients can access their dashboard by logging into the web-based platform to see a history of their assessments. Patient users complete a questionnaire that is based on the MDS-UPDRS. The questionnaire comprises 4 sections shown in Figure 1. Of these sections, 3 are multiple-choice questions based on daily living experiences; an example is shown in Figure 2, and there is also a motor examination where users perform tasks similar to those outlined in the MDS-UPDRS. Data are captured via 2 depth cameras (Intel models D435 and D455) that track a patient's movement in 3D. Before completing the motor examination, there is a series of ability questions that determine whether the user can safely perform all motor tasks; an example is shown in Figure 3. Motor tasks are explained in written form on the screen along with a demonstration video that presents users with a visual walk-through of the movement; an example is shown in Figure 4. The intelligent software scores each motor task based on the same parameters as the MDS-UPDRS. However, it is important to note that the system's scoring has not yet been validated. After users complete the 4 sections, a participant's responses are analyzed to put an individual on a PD rating scale from 0 to 4 (0=normal, 1=slight, 2=mild, 3=moderate, and 4=severe). A summary of a user's scores is presented on the home page, which can be seen in Figure 5. The survey was intentionally crafted and edited from the original MDS-UPDRS to use layperson language for easy comprehension. Although some technical terms appeared in titles or examples, they were not essential for answering questions or comprehending instructions.

Figure 1. The 4 sections of the Parkinson's Remote Interactive Monitoring System (PRIMS) questionnaire based on the Movement Disorder Society–sponsored revision of the Unified Parkinson's Disease Rating Scale.



XSL•FO
RenderX

Figure 2. Example multiple-choice question.

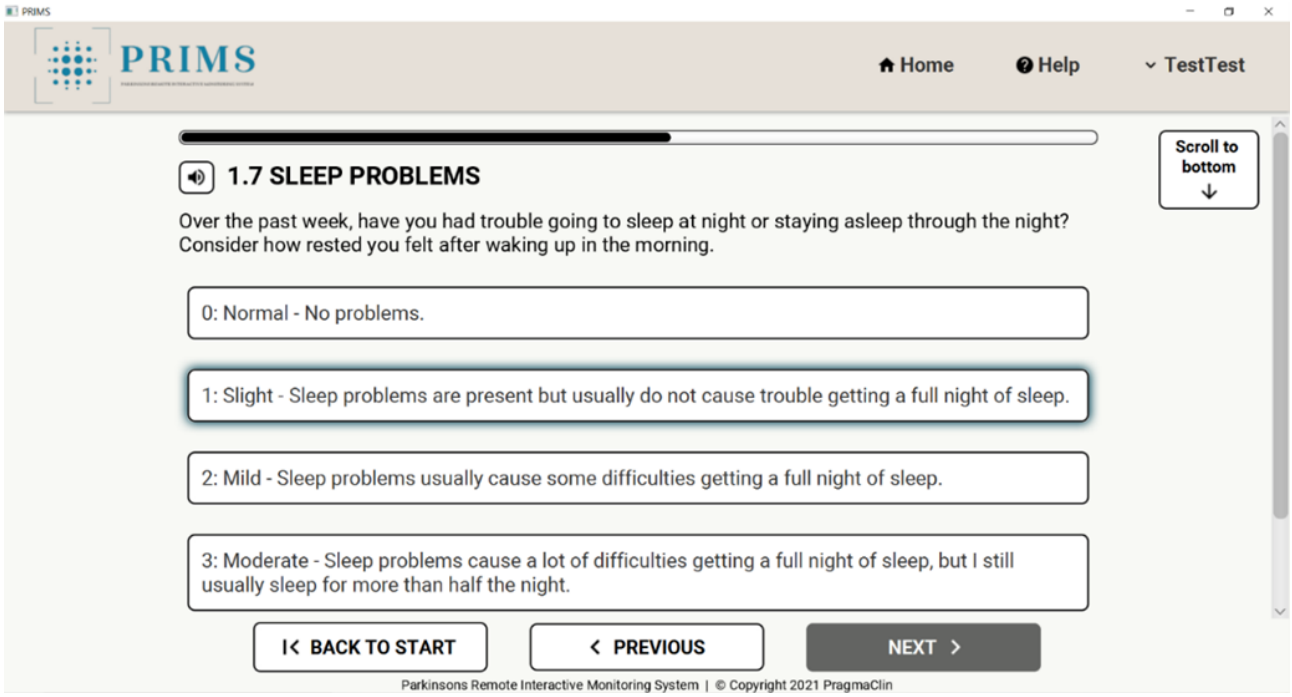
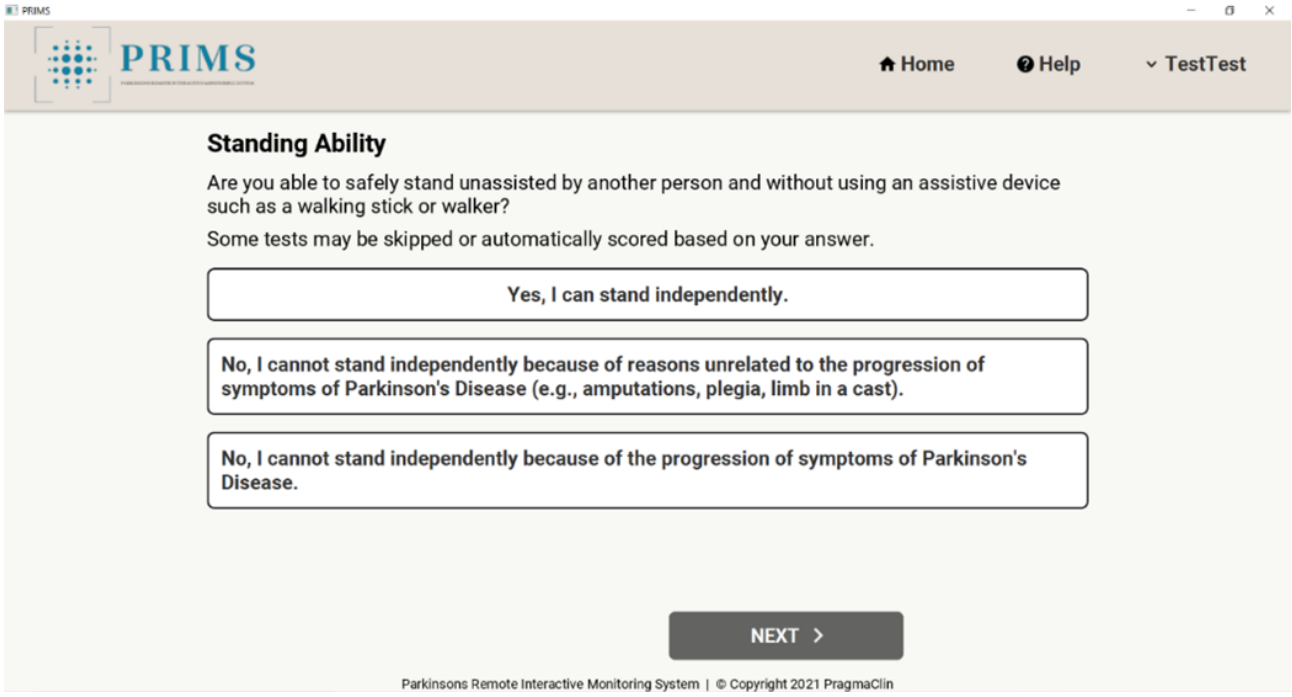


Figure 3. Multiple-choice question assessing an individual's ability to stand.



XSL•FO
RenderX

Figure 4. Example of a motor task page.

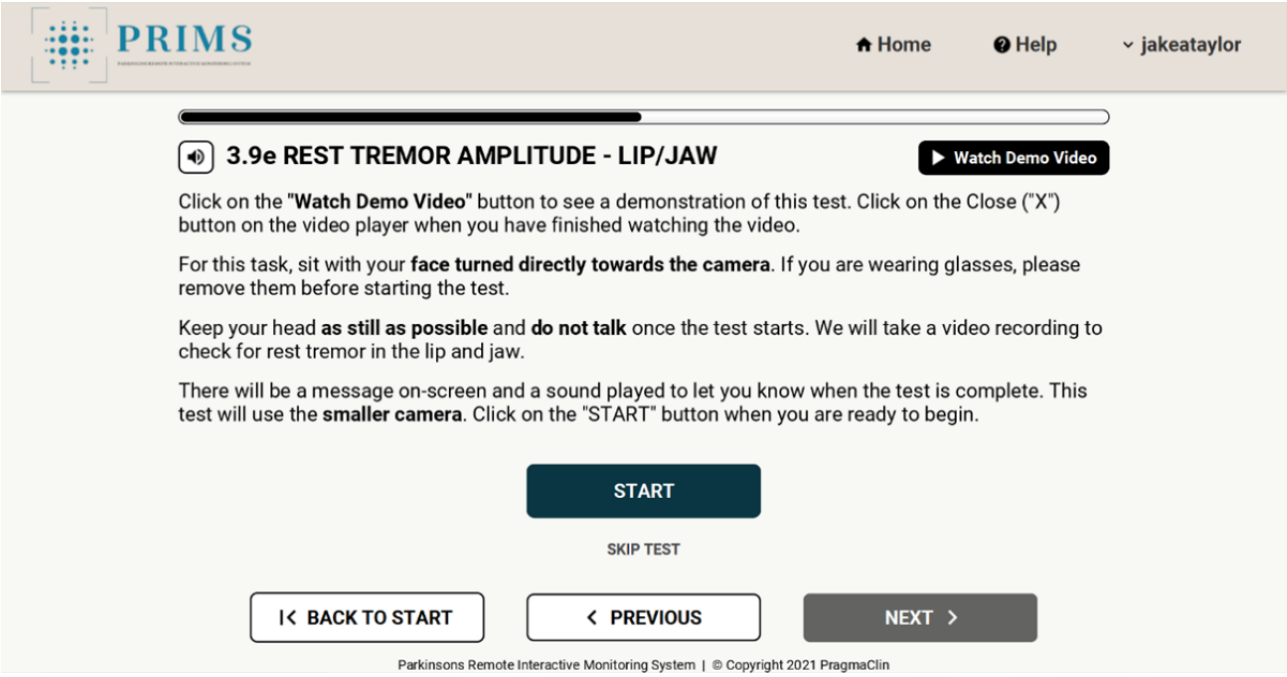
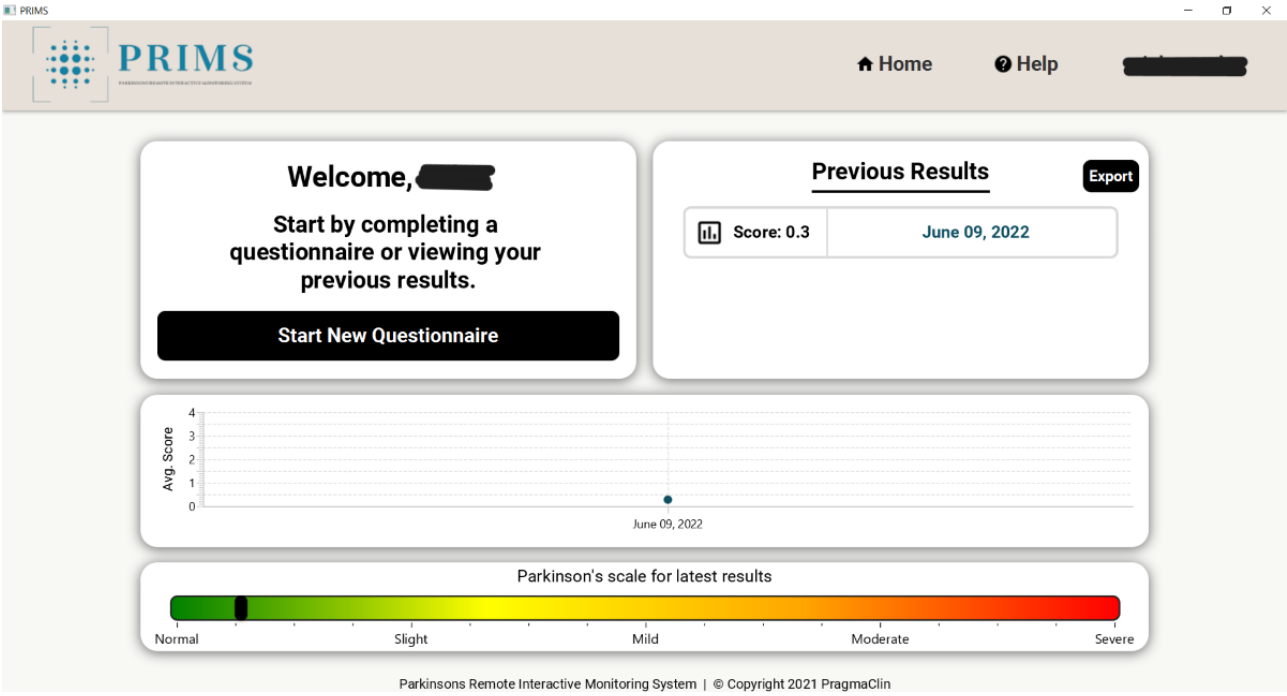


Figure 5. Screenshot of the Parkinson's Remote Interactive Monitoring System (PRIMS) home page.



Equipment

The PRIMS was run on a Dell G15 laptop computer, and Intel RealSense D435 (small—stand-alone mini tripod beside the laptop) and D455 (large—mounted on the computer) depth cameras were used. Participants sat on a contemporary midback task office chair with wheels for the entire questionnaire and had the option of using a Kensington Pro Fit wireless computer mouse. Interviews were recorded using a HyperX SoloCast stand-alone microphone. Audacity (Muse Group) was used as an audio recording and processing software. The computer-assisted qualitative coding software Delve (Twenty to Nine) was used for thematic analysis. All audio files were transcribed into Word (Microsoft Corp) before being uploaded to Delve. Qualtrics XM (Qualtrics International Inc) was used to administer the SUS survey and the virtual consent form.

Usability Testing Protocol and Procedure

Usability testing occurred at PragmaClin Research Inc's office site and was carried out by a trained research assistant (RA). Before the start of testing, the RA explained the study objective and research protocol to the participants. The RA also provided detailed information about the test procedures and described the purpose of the PRIMS. The example script is provided in Multimedia Appendix 1. As the research team was interested in how participants interact with the system when there is no

XSL•FO
RenderX

one present to assist them, the RA was not allowed to help unless deemed necessary. The necessity to intervene in the form of helpful hints or prompts or manipulating the system was operationally defined as any circumstance in which the participant was unable to progress through the system without aid. An observer was present during the entire session. All participants used the PRIMS only once and reported user experience from this single use.

After the participants gave informed consent, they were instructed to start using the PRIMS. The initial PRIMS developed by PragmaClin Research Inc was used during the usability testing. While participants worked their way through the assessment, they were encouraged to vocalize any confusion or ask any questions.

Data were recorded in the form of field notes, a short qualitative interview, and an SUS survey administered to participants after they completed the PRIMS questionnaire. Recorded data were used to identify a set of usability issues.

Field Notes

Structured observation was used to analyze the users' interactions with the system. During the session, the RA was instructed to observe and note any issues that arose along with any user critiques, comments, questions, difficulties, or observations about their interaction with the system. Thematic analysis was performed on these notes in Delve. Usability issues were identified via this method of analysis of the written notes.

Qualitative Interviews

After participants were finished with the PRIMS, they completed a short qualitative interview. The interview consisted of six questions: (1) What things did you like most about the PRIMS? (2) What things did you like least about the PRIMS? (3) Were there things about the PRIMS that you found confusing or frustrating? (4) What would you like to change about the PRIMS? (5) Are there any features that you would like to see added to the PRIMS? (6) Do you have any overall comments on the PRIMS?

Audio was transcribed and analyzed in the qualitative coding software Delve. Thematic analysis was performed following the framework by Braun and Clarke [53]. The themes were discussed, reviewed, and interpreted by the research team.

SUS Survey

After the interview was finished, participants completed a short-answer quantitative questionnaire following the standard SUS approach devised by Lewis and Sauro [54] in 2018. A copy of the survey is provided in Multimedia Appendix 2.

SUS scores were output in Qualtrics XM and then analyzed in Microsoft Excel. To calculate the SUS score, first the score contributions from each item (question) were summed. Each item's score contribution ranged from 0 to 4. For items 1, 3, 5, 7, and 9, the score contribution was the scale position minus 1. For items 2, 4, 6, 8, and 10, the contribution was 5 minus the scale position. We multiplied the sum of the scores by 2.5 to obtain the overall value of the SUS score. SUS scores have a range of 0 to 100. SUS scores of >70 points are considered acceptable usability (according to various other usability studies), and scores of >85 are regarded as excellent usability [55]. The curved grading scale by Lewis and Sauro [54] that was used to interpret the scores from the SUS is provided in Multimedia Appendix 3.

Results

Participants

A total of 11 people with PD participated in our study (female individuals: n=5, 45%; male individuals: n=6, 55%; age: mean 66.7, SD 7.77 years). Data from 91% (10/11) of the participants were fully analyzed as 1 user dropped out during the testing session. The 10 participants took, on average, 67.7 (SD 16.4) minutes to complete the motor examination and 84.2 (SD 23.3) minutes to complete the entire PRIMS questionnaire. Participants skipped 2.9 (SD 1.97) motor tests on average (total of 29 skipped tests). Table 1 shows which tests were skipped the most and how many times they were skipped.

XSL•FO
RenderX

Table 1. Number of times each motor test was skipped (tests ranked from the highest to lowest number of skips).

Motor test	Skipped tests, n
3.15	12
3.14	6
3.13	6
3.3	2
3.8	1
3.9e	1
3.10	1
3.1	0
3.2	0
3.4 and 3.5	0
3.6	0
3.7	0
3.9a-d	0
3.11	0
3.12	0

Field Notes

Thematic analysis of the observer's notes revealed 6 central usability problems associated with the PRIMS. These were the following: (1) automated voice prompts are confusing, (2) the small camera is problematic, (3) the motor test exhibits excessive sensitivity to the participant's orientation and position in relation to the cameras, (4) the system poses mobility challenges, (5) navigating the system is difficult, and (6) the motor test exhibits inconsistencies and technical issues.

Automated Voice Prompts Are Confusing

The RA noted on multiple occasions that the participants found the automated voice prompts to be confusing. Frequently, participants would begin the test and align themselves in a good position; however, when presented with audio prompts such as "make sure hand is not tilted" or "adjust hand position or angle," users would move in all directions:

> *The automated voice prompts were confusing and making it difficult. Even though P8 was turned in the right direction, the system still prompted him that he was turned the wrong way.*
>
> *A lot of automated verbal instructions were getting fired at P8, this made the task confusing and frustrating since they will be in the correct position and the software tells them wrong direction, turn right, etc.*

The prompt "Adjust hand position or angle" was especially frustrating and confusing to participants. Many verbalized their confusion with the statement. On multiple occasions, the RA noted that participants were becoming frustrated with the motor tests when the audio prompts began giving them instructions:

> *The prompt—no hand detected—is vague and confusing. P6 was in a good position, with their hand fully in view...Plus, P6 did not move and then immediately after there is a prompt saying—no hand detected...the audio prompts confused and frustrated P6.*

The RA also noted that some participants complained that the voice commands were authoritative and unfriendly:

> *...get in position and stand still—is a little authoritative! P8 mentioned this, participant did not like the automated voice instructions, said it needs to be more comforting / friendly. The automated voice prompts were authoritative P3 mentioned.*

Comments also often noted the repetitiveness of the automated voice prompts. They were repeated too often and in bizarre patterns, which caused confusion and frustration among users:

> *"don't move body in good position" was constantly repeating...was repeated ~5 times over, and the test wouldn't start.*
>
> *"Make sure hand is fully in view" is repeated even when the hand is fully in view, The system was prompting repeatedly "make sure hand is not tilted," motor test was very particular on positioning here. This made hand and body measurements very difficult for P2. They were moving their hand in all directions trying to figure out what tilted meant.*

Overall, participants found the automated voice prompts to be vague and irritating, as they rarely provided useful corrective feedback.

The Small Camera Is Problematic

The RA frequently noted problems associated with the tests that used the small camera or issues directly related to the small camera itself. The narrower field of view was an issue for

XSL•FO
RenderX

multiple tests; participants frequently moved out of the camera view midway through a task:

> *Finger to nose movements were difficult for P7, they had a hard time staying in the cameraview. They would be prompted that they are in a good position, then move out of the cameraview when performing their finger to nose movements. Hand moves out of frame...when the participant does hand rotations...small camera issue. P4 had great difficulty staying in the cameraview for these hand rotation tasks.*

The system would frequently tell users that they were in a good position, but when they started performing the task, they would move out of the camera view. The setup of the small camera also created problems for users. Users had to manipulate and adjust the small camera on the tripod to align themselves in a proper position. Frequently, this would lead to participants becoming uncomfortable due to the poor ergonomics of the system:

> *Face measurements using the small camera were not comfortable for P4. Small camera moved P4 into an uncomfortable position.*
>
> *The fact that P3 had to look at the camera for the test but look at the screen to get into position was causing difficulties.*
>
> *Getting into position was difficult for P9. Again, ergonomics is poor here for performing the test on two different sides when the camera is on one side.*

The tests that asked users to adjust the small camera were especially problematic. The RA frequently noted that people who pulled the cords out when laying the camera down for the final 2 tests did the following:

> *P2 moved the tripod/small camera, it was sloppy and difficult to work with. P2 pulled the cord out while adjusting the small camera.*
>
> *P11 unplugged the camera when they moved it for 3.15...issue with adjusting camera.*

Overall, the RA noted far more issues with the tests that used the small camera compared to those that used the larger camera.

The Motor Test Is Excessively Sensitive

It was clear from the RA's notes that participants had difficulty getting into what the system would consider to be valid positions to score during the motor examination. At times, the system would prompt users to stay still:

> *P9 had difficulty with hand measurements, they were unable to hold their hand still (issue since the software is for PwP). System is far too particular. The hand and body measurements require users to stay still to capture the measurement; P11 had great difficulty with this. System is far too particular on positioning here, which is just not feasible for those with PD.*

The system in its current state is very sensitive, which posed challenges for users:

> *The software is too picky on the positioning, participant's dyskinesia made it very difficult to stay in position, when the software told P7 that they were in a good position, they only had to move very slightly for the software to tell them that they needed to "adjust hand position or angle."*
>
> *A slight tilt is all it took for P8 to move out of position. The system was very particular on the positioning of the limb.*

The RA noted on multiple occasions that the system would repeat certain prompts when users were close to getting into the correct position:

> *"Don't move, hand is in good position" repeats a lot when P6 was "on the edge" of a good position. And when you start rotating your hand "no hand detected."*
>
> *"Don't move, hand is in good position" repeated a lot when P5 was "on the edge" of being in a good position.*
>
> *The system is far too sensitive, "Don't move face is in good position" kept repeating even though P3's face was in a good position. The motor test is too picky, its needs to be able to get the measurements from a broader range of places.*

Overall, it was clear from the recorded field notes that the motor examination was difficult for users:

> *3.14, and 3.15 P9 had a lot of difficulty keeping their hand in the correct position, P8 had difficulties getting their hand perfectly parallel to the camera face, The fact that P3 had to look at the camera for the test but look at the screen to get into position was causing difficulties.*

The System Poses Mobility Challenges

The setup of the PRIMS posed various mobility challenges for users. The RA noted on multiple occasions that users felt that the chair and constant movement of users were both issues. In its current state, the PRIMS requires participants to move back and forth from the computer to go from one task to another. The RA noted on several occasions that this posed a challenge for users:

> *Moving back and forth from the computer was difficult...the system requires too much movement of the chair, and to and from the computer, P7 vocalized this.*

This, coupled with the fact that users are constantly moving the chair in and out of the camera view, made the motor examination tiring for participants:

> *There is constant movement to and from the system that was tiring P8.*

The chair was difficult to work around due to the nature of the test; the RA noted repeatedly that participants failed to complete tests due to the chair obstructing the camera view:

XSL•FO
RenderX

> *...chair was in cameraview during these tests, posed an issue, was also difficult for participants to work around it.*

Some participants were also obstructed by the legs of the chair:

> *P5 needed to work around the chair legs for the arising from chair task, they vocalized this. The legs of the chair were in the way.*

Multiple participants also pointed out that they thought the chair was a safety issue:

> *The chair with wheels concerned P2. They thought it was a safety issue.*
>
> *P3 mentioned a few times that people with PD are told NOT to use chairs with wheels.*
>
> *For safety reasons, the chair with wheels is a huge problem, P7 vocalized this.*

Overall, the maneuvering required to complete the motor examination was an issue for users.

Navigating the System Is Difficult

There were frequent comments made by the RA on navigational issues users had while working through the system. Many participants had issues with the required amount of scrolling:

> *Scrolling is too difficult. a bigger screen would allow the entire survey to fit on one screen. P5 found the scrolling to be a challenge right away.*
>
> *P9 again demonstrated issues with scrolling and navigation, they had trouble scrolling to the bottom of the screen to select next on multiple tasks.*

Some users even found that they made mistakes due to the need to scroll to the bottom of the page:

> *P7 found that the scrolling led to mistakes. Choices they didn't mean to select.*

The RA also noted that the amount of clicking was an issue, specifically accurate clicking:

> *Skip test button is too small. Navigation issue, P9 had trouble clicking it due to dyskinesia, since it was so small.*
>
> *P7 had trouble closing the demo videos. Too much accurate clicking / total clicking required.*

Users also had issues with the computer mouse and mentioned that a touch screen interface would be preferred:

> *P9 had significant difficulties with the mouse. They said that they would prefer a touch screen.*
>
> *P2 didn't like using the mouse...Stated right away that they wanted to use a touch screen.*

The test window and demonstration video window also caused issues for users. Demonstration videos would frequently open in inconsistent sizes and locations of the screen, making them difficult to close:

> *Demo videos were opening in small windows at the top of the screen. Made them difficult to close and to watch.*

The test window did not show a married image of the person, which made it confusing to get into position:

> *Screen being non-mirrored is an issue. P5 had trouble moving into position because of this.*

The RA also noted that, as the software was not entirely full screen, users would frequently open other programs by accidentally clicking on the bottom task bar:

> *Full screen should eliminate the lower task bar (desktop), P4 ended up clicking things below or bringing up the news.*

Overall, users had difficulty navigating the system to progress.

The Motor Test Exhibits Technical Issues

There were frequent notes made on the system not operating correctly. Users would perform tests correctly or be told that they were in a good position yet would still be asked to try again as the system did not capture enough valid measurements to score:

> *Hand movements test stated that there were not enough valid measurements to score, after the participant did everything correctly.*

There were also occurrences in which users would perform tasks incorrectly and the test would still function:

> *Postural stability test worked even though the participant was not in the correct position at all.*
>
> *The foot tapping test ran even though P3 tapped the wrong foot, the system still gave him a score. They performed the measurement incorrectly, yet the system still considered it to be valid.*

This would lead to confusion among users as they would go through tests being told that they were in a good position without any other corrective feedback only to be asked to try again:

> *P6 performed the test correctly without any prompts to change position yet the system still prompted them to try again. The test will prompt people to start walking, and will run through without any corrective feedback, but may still state that there were not enough valid measurements to score.*

Some tests also tended to shut off very early and inconsistently. Other tests would often produce nonsense automated voice prompts:

> *...while performing the finger to nose movements: the audio prompt "multiple hands detected" was repeated even though there was only one hand in the camera view.*

Overall, the motor test presented frequent glitches causing usability trouble for participants.

Qualitative Interviews

Thematic analysis of qualitative interview responses revealed 4 central themes associated with participants' opinions on the PRIMS. These were (1) admiration of purpose, (2) excessive system sensitivity, (3) video instructions preferred, and (4) written instructions disliked.

XSL•FO
RenderX

Admiration of Purpose

Most of the participants showed an affinity for the overarching objective of the PRIMS:

> *I know what the main objective is, and I applaud that, that is a good objective.* [Participant 8]

They were excited about the system being available to people with PD:

> *I think it is awesome that people will have access to this.* [Participant 11]
>
> *I just like the fact that this is available for people.* [Participant 11]
>
> *I am sure a lot of people would be thrilled to have this at their doctor's office.* [Participant 11]

The intended purpose of the PRIMS was also well received and understood. Participants liked the idea that this will give their physicians a better view of their condition and support their ability to do their job:

> *Doctors often don't have a lot of time to do the examination in depth. My in-person examinations with my Neurologist are very fleeting, and scratching the surface in my view, but if that is the norm, and my Neurologist got a good reputation, then something like this would be very very helpful.* [Participant 8]
>
> *I like how then your doctor would have a better idea of what you are doing really, rather than based on that little scope of time kind of thing. Yea...that would be good.* [Participant 6]
>
> *I like that it can be used for long distance. And in our new post covid medical system, we need to free up time for our doctors.* [Participant 3]

Overall, participants admired the system and what it is trying to achieve and were excited to see the finished product in the future:

> *I think it would be worthwhile, if it was something that was worked out, if all the bugs and stuff were worked out it could be used as a tool...* [Participant 4]
>
> *Overall, I think it's a pretty good system. I think it will help patients or people.* [Participant 1]

Excessive System Sensitivity

Most participants found that the motor examination was very sensitive, which made it difficult to get into the proper position for the tests:

> *Yea like I say it's too sensitive, cause we have Parkinson's, and most people, you know you can be [shaky] and there's no way you are going to be able to stop it [tremors].* [Participant 2]

Users also found it frustrating and time-consuming:

> *Well I didn't like how it kept telling me that my hand is in a good position and then it's not or they don't detect it or those kinds of things, it can get a bit frustrating.* [Participant 1]
>
> *The only thing is the actual working of it in those couple of times where no matter what I did or didn't do, everything was as still as I could make it and it says you are fine then it says nope you have to start again...oh my gentle god...maybe it's just too sensitive or something.* [Participant 2]
>
> *It is time consuming too.* [Participant 11]
>
> *It is long and...you know...as Parkinson's patients you get tired easily.* [Participant 6]

The frustration seemed to stem from the fact that the system was very particular on how it wanted users to be positioned. Participants had the greatest trouble with the hand movements and tasks that required users to stay still:

> *The ones we have the most difficulty with are the hand.* [Participant 8]
>
> *Because holding still is a challenge for some people with Parkinson's...some people have tremor, and some don't. For those that do, holding still is a real challenge.* [Participant 11]

Overall, users found that the system was difficult to use in its current state due to its sensitivity:

> *I think that that [PRIMS] would be difficult for some people...unless you had extensive training.* [Participant 4]

Video Instructions Preferred

Users found that the demonstration videos were far more helpful, and much less confusing, than the written instructions:

> *The video was a good tool because we have a lot of brain fog, and reading can be confusing and looking at the video makes things much easier.* [Participant 3]
>
> *...you are able to see a video of the man actually doing what you are supposed to do you know is quite helpful too I thought.* [Participant 8]
>
> *There are a lot of words there, in the instructions, again I think if you had it in bullet form maybe. It's a lot easier to watch the video.* [Participant 1]
>
> *The video was good to show how to do the testing.* [Participant 1]

Users suggested more video instructions and less written instructions and even suggested an introduction video outlining what the system entails:

> *Instead of just jumping right in there, if you had, well I guess it would be a video, but if you had a synopsis of what the testing involved. Maybe if we had a 10-minute video overlooking the whole test at first.* [Participant 1]

Overall, the videos were one of the most liked aspects of the entire system:

> *The things that I thought worked best were the videos.* [Participant 5]
>
> *I like how there is a video...watching what they do is much more clear.* [Participant 11]

XSL•FO
RenderX

Written Instructions Disliked

Many participants found that the written instructions were vague and confusing:

> *Some of the tests I found confusing, but again that was the written instructions that were somewhat confusing.* [Participant 5]
>
> *The instructions were really kind of vague.* [Participant 1]

Multiple users stated that these instructions were annoying and far too wordy:

> *Too much instruction, yea, but I know you have to have the instruction down, but it was a lot of reading.* [Participant 1]
>
> *I would say all the text that open on the screen, yea it's like going to a presentation...mostly just the instructions, I mean you're asking me, in my gut, kind of what I found to be annoying about it...and...the text was annoying.* [Participant 7]

Participants suggested that more concise bullet-point instructions would be preferred over written paragraphs:

> *I think the instructions was too many...If it was concise and shorter instructions, I think it would make it a little better.* [Participant 1]

SUS Survey

The average SUS score was calculated to be 69.2, which corresponds to a C on our curved grading scale [54]. The PRIMS failed to meet the acceptable SUS score of 70.

Discussion

Principal Findings

We conducted a multiple methods study to assess the usability, functionality, and user experience of the PRIMS. Thematic analysis of interview transcripts and field notes revealed multiple themes and usability issues, respectively, that describe the tested product. An SUS survey also gave us a key objective insight into the system and its user experience.

One of the key findings of this study was that video instructions were preferred over written instructions. Thematic analysis of interview transcripts revealed these 2 themes (*written instructions disliked* and *video instructions preferred*). Multiple participants stated that the video instructions were much less confusing and much more informative than the on-screen text. The written instructions were designed to give all the necessary information to complete the task. This may have resulted in users feeling unmotivated to read the entire set of instructions as there was an intimidating amount of text present on-screen. Other investigations comparing video to written instructions have found similar results. Cosford et al [56] evaluated the effectiveness of video and handout instructions during a veterinary student examination. Their findings revealed that students using video instructions achieved notably higher scores, suggesting a better understanding of the tasks compared to those using handouts [56]. Shah and Gupta [57] found that video instructions were significantly more effective than written instructions in teaching inhaler use technique. Video instructions provide both a visual and audio description of each task, which can make the instructions both clearer and less time-consuming.

Another principal finding was that the PRIMS motor examination was too sensitive and particular on users' body positions during the tests. Thematic analysis of field notes and interview transcripts unveiled 2 areas of issues, namely, system sensitivity and the motor test's positioning specificity, which exhibited alignment in their respective scopes. To quantify each motor task performed during the PRIMS questionnaire, the depth cameras would require participants to be oriented in a "good position." From the RA's observational notes and the interview transcripts, it was clear that the system asked too much of users, which led to frustration and difficulties. Systems designed for those with movement disorders must be accommodating to their needs. The PRIMS, in its current state, asks users to stay still in certain situations and adopt specific and uncomfortable positions to score their movements. Future versions of the PRIMS will need to address this in their design and implementation.

Another theme revealed from analysis of field notes was that the automated voice prompts that are used during each motor test are confusing to participants (*automated voice prompts are confusing*). The prompts would also do more harm than good when it came to helping participants align themselves in the correct position for each motor test. It was noted that users tended to move in all directions in response to the automated voice. This could be due to the vague nature of the instructions provided by the prompts. They also led to frustration and confusion, making them an ineffective tool to guide users through the tests. Some users even stated that they found the automated voice to be authoritative and unfriendly, which only increased their frustration with the system. The consensus of this key finding was that these automated prompts did not provide any useful corrective feedback and only led to confusion and frustration among participants. Mays et al [58] delved into how people in the United States perceive automated communication, such as interactive voice response systems. They found that older respondents especially did not enjoy the automated voice system and exhibited greater levels of frustration toward it [58]. Most people with PD are older individuals; therefore, it would be best practice to tailor the system's instructions and prompts to their typical preferences.

Our next key finding is that the small camera tended to cause more problems for users than the large camera. Thematic analysis performed on field notes revealed this theme (*the small camera is problematic*). The smaller-depth camera (Intel D435) has a narrower field of view and was primarily used for the hand movement tests during the motor examination. A common problem that users faced was staying within the camera view for these tests. As the RA noted on several occasions, it was common for participants to have difficulties with the narrow view. The position of the small camera also caused trouble for users. Unlike the larger camera, the smaller camera is placed on a tripod on either side of the computer (Figure 6). It was noted frequently in the field notes that users had to manipulate and adjust this camera to align themselves in the proper position. The placement of the camera also negatively impacted the

XSL•FO
RenderX

system's ergonomics. In addition, for tests 3.14 and 3.15, users were prompted to flip the tripod down so that the camera was facing the ceiling. Frequently, users accidentally disconnected the cords from the camera when moving it. In general, a setup in which users do not have to adjust any equipment would be preferential. To our knowledge, there is no direct study to compare this finding to.

Figure 6. Diagram of hardware setup showing the position of the laptop and 2 depth cameras.

Another principal finding of our thematic analysis was that the PRIMS has associated mobility challenges for users. A big issue that the design of the system has is the constant movement to and from the computer to go from one task to another. At times when the users' feet or whole body had to be visible to the camera, participants would have to move backward until this was the case. To move onto the next task, users would have to return to the laptop to select it. This way, users are constantly moving to and from the computer and frequently being obstructed by the chair. Several users even pointed out that they felt as though others would have problems with the back-and-forth nature of the system. When designing systems for those with movement disorders, it is important to consider the user experience as a whole.

Thematic analysis of field notes also revealed that users had difficulties navigating the system (*navigating the system is difficult*). In its tested state, the PRIMS was operated using a standard computer mouse or laptop touchpad (depending on user preference). Navigating the system using either of these tools caused difficulty among participants. Scrolling or clicking to move in between sections of the questionnaire was frequently noted as a challenge for users and even led to mistakes in some circumstances. The accuracy of the clicking to move through the system was also a big issue and would often lead to opening other applications or closing the PRIMS software. It is important to remember that, when designing systems for those with movement disorders, there must be special considerations taken. A viable option could be a touch screen device featuring prominently sized buttons, eliminating the need for scrolling. However, it is worth noting that touch screens can lead to increased postural discomfort during use [59]. Thus, offering a variety of system operation methods might be the most effective way to cater to diverse user needs and preferences. Enhancing usability is paramount not only for optimizing human-computer interactions but also for ensuring the system's social and practical acceptance [48]. A system's usability should be of a standard that facilitates effortless task execution by the user. Given that the PRIMS posed challenges for users in performing certain tasks, there is a clear need for usability enhancements.

Another principal finding that came from analyzing the field notes is that the motor examination did not function perfectly or as intended. Similar to any new software system, the PRIMS had its share of technical issues. One of the most frequent issues noted by the RA was that users would perform tests correctly yet the system would fail to score their movements. This is a problem especially when the system does not prompt any corrective feedback yet still informs users that there were not enough valid measurements to score. The software issues caused frustration among the participants and led to multiple usability issues. Usability is tied to functionality, although they are not exactly the same. When a product is not functioning correctly, it ultimately impacts its usability. Thus, ensuring that the system works as it is intended must be a priority for future developers to improve its usability.

Through the analysis of interview transcripts, a prominent theme of appreciation emerged in relation to the PRIMS. Designed with the primary objective of enhancing care for individuals living with PD, the PRIMS was met with significant enthusiasm. Users recognized the immense potential of such a remote monitoring solution, expressing eagerness about its availability. This positive reception underscores the importance of aligning the product's design with the needs of its intended users. The participants' commendation of both the system and its mission suggests that the overarching principle of the product is robust. Previous studies have emphasized the pivotal role of consumer perspectives in determining the success of a product [60]. Given this context, such a positive reception indicates a promising trajectory for the future deployment of the PRIMS.

Our last principal finding from this usability study was that our recorded SUS score was 69.7, which failed to meet our acceptable usability score of 70. There are several reasons that

XSL•FO
RenderX

could explain why participants felt that the system was not as user-friendly as it should be. Binyamin et al [61] used the SUS to evaluate a learning management system in an educational setting. Their system also failed to meet an acceptable usability score of 70. A distinct aspect of their study was that participants engaged with the system repeatedly throughout a summer term. They observed a direct relationship between the frequency of system use and the SUS score, suggesting that increased familiarity led to improved usability ratings [61]. Drawing from this, it is conceivable that, if participants in our study had interacted with the system over an extended duration, as intended for the PRIMS, the SUS score might have been more favorable due to enhanced user familiarity by the study's conclusion.

Severity of Usability Problems

Within the spectrum of presented usability problems, a hierarchical assessment of severity becomes imperative considering factors ranging from potential risks to participant safety to issues causing minor hindrances in task completion. We ranked the issues posing threats to safety and mobility as top priorities to be addressed, followed by those that led to difficulty and frustration, with minor issues that may have slowed participants down being of the least concern.

Foremost among the identified challenges were those associated with mobility constraints, standing out as the most severe due to their inhibiting impact on participants with mobility challenges. Beyond impeding the use of the PRIMS, these challenges pose a risk to participant safety by potentially placing individuals in vulnerable positions. Notably, the constant need for movement to and from the computer for task transitions emerges as a top priority for resolution.

A usability problem of a lesser degree of severity pertains to the system's sensitivity, specifically in quantifying motor tasks during the PRIMS questionnaire. The requirement for participants to be consistently in a "good position" proved overly demanding, leading to frustration and difficulties, as observed in RA notes and interview transcripts. This underscores the importance of designing systems for individuals with movement disorders to be accommodating to their unique needs. The current state of the PRIMS, requiring users to stay still in certain situations and adopt uncomfortable positions, resulted in skipped tests, increasing the priority of addressing this issue.

Issues that made completion difficult included challenges in navigating using the computer mouse and occasional malfunctions in the motor examination. These concerns, while not as severe as mobility-related issues or the system's sensitivity, warrant attention as they contribute to user frustration and impact task completion.

Finally, minor difficulties associated with operating the small camera and managing automated voice prompts and written instructions require fine refinements rather than constituting significant hurdles. While not impeding overall system navigation, these issues contributed to slower completion and user frustration. In the hierarchy of severity, they represent areas for enhancement rather than critical concerns requiring immediate attention.

Implications

This study conducted a comprehensive evaluation of a system specifically tailored for individuals with motor and cognitive conditions, shedding light on critical considerations for the development of technology for this population. First, we advise against the use of desktop applications requiring a computer mouse, scrolling, and intricate clicking, recognizing the potential challenges faced by users. Furthermore, our findings emphasize the superiority of visual instructions over written ones, also suggesting that automated voice prompts should be used judiciously and presented in a friendly manner and offer clear instructions, especially during confirmation processes. To address the mobility challenges commonly faced by this population, systems necessitating movements in front of a camera should minimize the need for multiple adjustments as these can introduce errors. In addition, our study underscores the importance of system flexibility, allowing for a significant margin of error in data capture without imposing the requirement for participants to remain perfectly still during calibration—an often-unattainable feat for those with motor conditions. As we navigate future developments of the PRIMS, these insights will serve as a guide in creating a product that effectively addresses the outlined issues, emphasizing a visually guided interface requiring minimal effort for seamless operation, aligning with the unique needs of our target user base.

Limitations and Future Investigation

There were several limitations to this study. Our small sample size may not have revealed all the usability issues [62,63] as testing with a small number of participants tends to only reveal the major flaws or glitches in the system. However, our main objective was to uncover the biggest areas of concern rather than identifying every problem associated with the system.

Another limitation to this study was our methodology. Other common qualitative data recording techniques for usability studies include the think-aloud technique [64] and focus groups. The think-aloud technique is the process in which users are encouraged to verbalize their perceptions as they interact with the system [64], which can provide insight into the user experience. Focus groups with study participants following the interviews could have produced richer information. This would have given users the opportunity to compare their ideas and thoughts on the PRIMS. However, while conducting this study, we made a deliberate decision not to use the think-aloud technique. This choice was grounded in our consideration for the unique challenges faced by individuals with PD, particularly those with motor and speech difficulties. The think-aloud technique traditionally involves participants verbalizing their thoughts as they navigate through a system. However, given the potential speech impediments, tremors, and other motor-related challenges associated with PD, we anticipated that asking participants to vocalize their thoughts could introduce unnecessary stress and frustration, and we did not want to pile on any extra cognitive load. To ensure a more comfortable and authentic testing environment for individuals with PD, we opted for direct observation followed by an interview after they were finished using the system, allowing

XSL•FO
RenderX

us to carefully note any challenges users encountered as they interacted with the software.

We also acknowledge that other methods, for example, mixed methods research [65], are applicable for usability and user experience research. Mixed methods may allow for a more comprehensive understanding of a user's experience, which can enable researchers to identify specific usability issues [66]. There are also other scales that we could have used to quantify user satisfaction, for example, the Post-Study System Usability Questionnaire [67]. As our product is in the early stages of development, we opted for a simpler multiple methods study to uncover the major flaws in our system. Future usability studies on the PRIMS can use a mixed methods design to gain a deeper understanding of usability issues.

Considering the intricacies involved in designing systems for users with movement disorders, a promising direction for future research could entail conducting user-centered design studies to tackle the identified usability challenges. This approach, which is advocated by other authors as an effective methodology for achieving a usable product, aims to design products that consider the needs and interests of end users [68-71]. Salinas et al [72] have reviewed the techniques and tools used in the successful redesigns of graphical user interfaces of software products following the user-centered design approach. While some of these techniques align with those used in our evaluation, the key lies in using these techniques throughout the design process instead of solely during product testing. Commonly reported methods of user testing include prototyping, pre- and postdesign interviews, heuristic evaluation, and surveys or questionnaires [72]. Therefore, upcoming research endeavors concerning the redesigned PRIMS should embrace a user-centered design methodology to guarantee the satisfaction of end users' needs and explore the integration of the aforementioned effective techniques. Moreover, future investigations could concentrate on crafting customizable interface options enabling users to tailor their interaction experience according to their individual capabilities and preferences. This might entail the incorporation of adjustable settings for font size, button layout, and navigation pathways such as voice activation or remote controllers to cater to a diverse range of users.

The current iteration of the PRIMS faces practicality challenges for home use. A more feasible adaptation would necessitate enhanced usability, reduced equipment costs, and minimal space requirements. Substantial updates are imperative to transform the PRIMS into a valuable home-based tool. This entails enhancing user-friendliness, optimizing the product to function seamlessly on common smartphone or tablet cameras, and refining the interface for a user-friendly experience.

It is important to emphasize that a direct score comparison between the PRIMS and a clinician was not conducted in this study. The PRIMS did assign scores to each movement in the motor test using an algorithm developed by PragmaClin Research Inc. However, it is essential to clarify that this study exclusively focused on usability and did not assess the validity of the scoring process. The validation of scoring algorithms remains a subject for future investigations.

Conclusions

In conclusion, the PRIMS currently exhibits several usability challenges that hinder its efficient use by individuals with PD. For the system to achieve successful implementation and gain broad acceptance, it is imperative to address these identified issues. Feedback from this study is being used to upgrade the PRIMS so that it better aligns with patients' needs. This study contributes significantly to the growing literature on usability testing, particularly emphasizing design nuances for systems tailored to those with movement disorders. Moving forward, it would be beneficial for future research to explore diverse interaction methods with digital devices, aiming to pinpoint optimal usability practices.

Acknowledgments

The authors would like to thank the Parkinson Society Newfoundland and Labrador for assistance with recruitment and the Mitacs Accelerate program for funding this project. They also extend their heartfelt gratitude to the participants, whose invaluable contributions have been instrumental in shaping this research endeavor. This study is one of several manuscripts that will collectively contribute to BB's PhD thesis.

Data Availability

The data sets generated during and analyzed during this study are available in the Memorial University Dataverse-Borealis repository [73].

Conflicts of Interest

BB and JT are employees of PragmaClin Research Inc, the company that developed the Parkinson's Remote Interactive Monitoring System. All other authors declare no other conflicts of interest.

Multimedia Appendix 1
Usability study script.
[PDF File (Adobe PDF File), 142 KB - humanfactors_v11i1e54145_app1.pdf]

XSL•FO
RenderX

Multimedia Appendix 2
System usability survey.
[PDF File (Adobe PDF File), 46 KB - humanfactors_v11i1e54145_app2.pdf]

Multimedia Appendix 3
Curved grading scale by Lewis and Sauro [55].
[PDF File (Adobe PDF File), 96 KB - humanfactors_v11i1e54145_app3.pdf]

References

1. Dorsey ER, Sherer T, Okun MS, Bloem BR. The emerging evidence of the Parkinson pandemic. J Parkinsons Dis 2018 Dec 18;8(s1):S3-S8. [doi: 10.3233/jpd-181474]
2. Ball N, Teo WP, Chandra S, Chapman J. Parkinson's disease and the environment. Front Neurol 2019 Mar 19;10:218 [FREE Full text] [doi: 10.3389/fneur.2019.00218] [Medline: 30941085]
3. Marras C, Beck JC, Bower JH, Roberts E, Ritz B, Ross GW, et al. Prevalence of Parkinson's disease across North America. NPJ Parkinsons Dis 2018 Jul 10;4:21 [FREE Full text] [doi: 10.1038/s41531-018-0058-0] [Medline: 30003140]
4. Wijeratne T, Grisold W, Trenkwalder C, Carroll W. World Brain Day 2020: move together to end Parkinson's disease. Lancet Neurol 2020 Aug;19(8):643. [doi: 10.1016/S1474-4422(20)30215-5] [Medline: 32702326]
5. Davie CA. A review of Parkinson's disease. Br Med Bull 2008;86:109-127. [doi: 10.1093/bmb/ldn013] [Medline: 18398010]
6. Cacabelos R. Parkinson's disease: from pathogenesis to pharmacogenomics. Int J Mol Sci 2017 Mar 04;18(3):551 [FREE Full text] [doi: 10.3390/ijms18030551] [Medline: 28273839]
7. Antonini A, Barone P, Marconi R, Morgante L, Zappulla S, Pontieri FE, et al. The progression of non-motor symptoms in Parkinson's disease and their contribution to motor disability and quality of life. J Neurol 2012 Dec;259(12):2621-2631. [doi: 10.1007/s00415-012-6557-8] [Medline: 22711157]
8. Statistics. Parkinson's Foundation. URL: https://www.parkinson.org/Understanding-Parkinsons/Statistics [accessed 2024-05-08]
9. Busis NA, Shanafelt TD, Keran CM, Levin KH, Schwarz HB, Molano JR, et al. Author response: burnout, career satisfaction, and well-being among US neurologists in 2016. Neurology 2017 Oct 10;89(15):1650-1651. [doi: 10.1212/WNL.0000000000004484] [Medline: 28993531]
10. Dall TM, Storm MV, Chakrabarti R, Drogan O, Keran CM, Donofrio PD, et al. Supply and demand analysis of the current and future US neurology workforce. Neurology 2013 Jul 30;81(5):470-478 [FREE Full text] [doi: 10.1212/WNL.0b013e318294b1cf] [Medline: 23596071]
11. Schapira AH, Obeso J. Timing of treatment initiation in Parkinson's disease: a need for reappraisal? Ann Neurol 2006 Mar;59(3):559-562. [doi: 10.1002/ana.20789] [Medline: 16489611]
12. Caslake R, Taylor KS, Counsell CE. Parkinson's disease misdiagnosed as stroke. BMJ Case Rep 2009;2009:bcr07.2008.0558 [FREE Full text] [doi: 10.1136/bcr.07.2008.0558] [Medline: 21686747]
13. DeMaagd G, Philip A. Parkinson's disease and its management: part 1: disease entity, risk factors, pathophysiology, clinical presentation, and diagnosis. P T 2015 Aug;40(8):504-532 [FREE Full text] [Medline: 26236139]
14. Morgan J, Sethi KD. Levodopa and the progression of Parkinson's disease. Curr Neurol Neurosci Rep 2005 Jul;5(4):261-262. [doi: 10.1007/s11910-005-0069-1] [Medline: 15987608]
15. Papapetropoulos SS. Patient diaries as a clinical endpoint in Parkinson's disease clinical trials. CNS Neurosci Ther 2012 May;18(5):380-387 [FREE Full text] [doi: 10.1111/j.1755-5949.2011.00253.x] [Medline: 22070400]
16. Martinez-Martin P, Rodriguez-Blazquez C, Forjaz MJ, Chaudhuri KR. Motor impairment and disability scales. In: Guide to Assessment Scales in Parkinson’s Disease. Tarporley: Springer Healthcare; Oct 17, 2014.
17. Ramaker C, Marinus J, Stiggelbout AM, Van Hilten BJ. Systematic evaluation of rating scales for impairment and disability in Parkinson's disease. Mov Disord 2002 Sep;17(5):867-876. [doi: 10.1002/mds.10248] [Medline: 12360535]
18. Goetz CG, Tilley BC, Shaftman SR, Stebbins GT, Fahn S, Martinez-Martin P, et al. Movement Disorder Society-sponsored revision of the Unified Parkinson's Disease Rating Scale (MDS-UPDRS): scale presentation and clinimetric testing results. Mov Disord 2008 Nov 15;23(15):2129-2170. [doi: 10.1002/mds.22340] [Medline: 19025984]
19. Bot BM, Suver C, Neto EC, Kellen M, Klein A, Bare C, et al. The mPower study, Parkinson disease mobile data collected using ResearchKit. Sci Data 2016 Mar 03;3:160011 [FREE Full text] [doi: 10.1038/sdata.2016.11] [Medline: 26938265]
20. Ossig C, Antonini A, Buhmann C, Classen J, Csoti I, Falkenburger B, et al. Wearable sensor-based objective assessment of motor symptoms in Parkinson's disease. J Neural Transm (Vienna) 2016 Jan;123(1):57-64. [doi: 10.1007/s00702-015-1439-8] [Medline: 26253901]
21. Silva de Lima AL, Hahn T, de Vries NM, Cohen E, Bataille L, Little MA, et al. Large-scale wearable sensor deployment in Parkinson's patients: the Parkinson@Home study protocol. JMIR Res Protoc 2016 Aug 26;5(3):e172 [FREE Full text] [doi: 10.2196/resprot.5990] [Medline: 27565186]

XSL•FO
RenderX

22. Post B, Merkus MP, de Bie RM, de Haan RJ, Speelman JD. Unified Parkinson's disease rating scale motor examination: are ratings of nurses, residents in neurology, and movement disorders specialists interchangeable? Mov Disord 2005 Dec;20(12):1577-1584. [doi: 10.1002/mds.20640] [Medline: 16116612]
23. Palmer JL, Coats MA, Roe CM, Hanko SM, Xiong C, Morris JC. Unified Parkinson's disease rating scale-motor exam: inter-rater reliability of advanced practice nurse and neurologist assessments. J Adv Nurs 2010 Jun;66(6):1382-1387 [FREE Full text] [doi: 10.1111/j.1365-2648.2010.05313.x] [Medline: 20546368]
24. Bhidayasiri R, Martinez-Martin P. Clinical assessments in Parkinson's disease: scales and monitoring. Int Rev Neurobiol 2017;132:129-182. [doi: 10.1016/bs.irn.2017.01.001] [Medline: 28554406]
25. Weintraub D, Comella CL, Horn S. Parkinson's disease--part 1: pathophysiology, symptoms, burden, diagnosis, and assessment. Am J Manag Care 2008 Mar;14(2 Suppl):S40-S48 [FREE Full text] [Medline: 18402507]
26. Larson DN, Schneider RB, Simuni T. A new era: the growth of video-based visits for remote management of persons with Parkinson's disease. J Parkinsons Dis 2021;11(s1):S27-S34 [FREE Full text] [doi: 10.3233/JPD-202381] [Medline: 33492246]
27. Dorsey ER, Achey MA, Beck CA, Beran DB, Biglan KM, Boyd CM, et al. National randomized controlled trial of virtual house calls for people with Parkinson's disease: interest and barriers. Telemed J E Health 2016 Jul;22(7):590-598 [FREE Full text] [doi: 10.1089/tmj.2015.0191] [Medline: 26886406]
28. Adams JL, Myers TL, Waddell EM, Spear KL, Schneider RB. Telemedicine: a valuable tool in neurodegenerative diseases. Curr Geriatr Rep 2020 Jun;9(2):72-81 [FREE Full text] [doi: 10.1007/s13670-020-00311-z] [Medline: 32509504]
29. Dorsey ER, Wagner JD, Bull MT, Rizzieri A, Grischkan J, Achey MA, et al. Feasibility of virtual research visits in fox trial finder. J Parkinsons Dis 2015;5(3):505-515 [FREE Full text] [doi: 10.3233/JPD-150549] [Medline: 26406130]
30. Espay AJ, Bonato P, Nahab FB, Maetzler W, Dean JM, Klucken J, et al. Technology in Parkinson's disease: challenges and opportunities. Mov Disord 2016 Sep;31(9):1272-1282 [FREE Full text] [doi: 10.1002/mds.26642] [Medline: 27125836]
31. Rovini E, Maremmani C, Cavallo F. How wearable sensors can support Parkinson's disease diagnosis and treatment: a systematic review. Front Neurosci 2017 Oct 06;11:555 [FREE Full text] [doi: 10.3389/fnins.2017.00555] [Medline: 29056899]
32. Dewey DC, Miocinovic S, Bernstein I, Khemani P, Dewey RB3, Querry R, et al. Automated gait and balance parameters diagnose and correlate with severity in Parkinson disease. J Neurol Sci 2014 Oct 15;345(1-2):131-138 [FREE Full text] [doi: 10.1016/j.jns.2014.07.026] [Medline: 25082782]
33. Wang J, Gong D, Luo H, Zhang W, Zhang L, Zhang H, et al. Measurement of step angle for quantifying the gait impairment of Parkinson's disease by wearable sensors: controlled study. JMIR Mhealth Uhealth 2020 Mar 20;8(3):e16650 [FREE Full text] [doi: 10.2196/16650] [Medline: 32196458]
34. Anand V, Bilal E, Ramos V, Naylor M, Demanuele C, Zhang H, et al. F62. Automatic detection of ON/OFF states in Parkinson disease patients using wearable inertial sensors. Clin Neurophysiol 2018 May;129(Supplement 1):e90. [doi: 10.1016/j.clinph.2018.04.225]
35. Lipsmeier F, Simillion C, Bamdadian A, Tortelli R, Byrne LM, Zhang YP, et al. A remote digital monitoring platform to assess cognitive and motor symptoms in Huntington disease: cross-sectional validation study. J Med Internet Res 2022 Jun 28;24(6):e32997 [FREE Full text] [doi: 10.2196/32997] [Medline: 35763342]
36. Joshi R, Bronstein JM, Keener A, Alcazar J, Yang DD, Joshi M, et al. PKG movement recording system use shows promise in routine clinical care of patients with Parkinson's disease. Front Neurol 2019 Oct 1;10:1027 [FREE Full text] [doi: 10.3389/fneur.2019.01027] [Medline: 31632333]
37. Chaudhuri KR, Hand A, Obam F, Belsey J. Cost-effectiveness analysis of the Parkinson's KinetiGraph and clinical assessment in the management of Parkinson's disease. J Med Econ 2022;25(1):774-782 [FREE Full text] [doi: 10.1080/13696998.2022.2080437] [Medline: 35593687]
38. StrivePD homepage. StrivePD. URL: https://www.strive.group/ [accessed 2024-05-08]
39. Wearable sensors - APDM wearable technologies. APDM Wearable Technologies. URL: https://apdm.com/wearable-sensors/ [accessed 2024-05-02]
40. Ramos V, Naylor M, Demanuele C, Zhang H, Amato S, Wacnik P, et al. T153. Wearable inertial sensors produce reliable endpoints in healthy volunteers and detect levodopa -induced changes in Parkinson’s disease patients. Clin Neurophysiol 2018 May;129(Supplement 1):e61. [doi: 10.1016/j.clinph.2018.04.154]
41. Piau A, Wild K, Mattek N, Kaye J. Current state of digital biomarker technologies for real-life, home-based monitoring of cognitive function for mild cognitive impairment to mild Alzheimer disease and implications for clinical care: systematic review. J Med Internet Res 2019 Aug 30;21(8):e12785 [FREE Full text] [doi: 10.2196/12785] [Medline: 31471958]
42. Pasluosta CF, Gassner H, Winkler J, Klucken J, Eskofier BM. An emerging era in the management of Parkinson's disease: wearable technologies and the internet of things. IEEE J Biomed Health Inform 2015 Nov;19(6):1873-1881. [doi: 10.1109/JBHI.2015.2461555] [Medline: 26241979]
43. Hergenroeder AL, Barone Gibbs B, Kotlarczyk MP, Kowalsky RJ, Perera S, Brach JS. Accuracy of objective physical activity monitors in measuring steps in older adults. Gerontol Geriatr Med 2018 Jun 17;4:2333721418781126 [FREE Full text] [doi: 10.1177/2333721418781126] [Medline: 29977979]
44. Heldman DA, Giuffrida JP, Cubo E. Wearable sensors for advanced therapy referral in Parkinson's disease. J Parkinsons Dis 2016 Jul 02;6(3):631-638. [doi: 10.3233/JPD-160830] [Medline: 27392872]

XSL•FO
RenderX

45. Ramdhani RA, Khojandi A, Shylo O, Kopell BH. Optimizing clinical assessments in Parkinson's disease through the use of wearable sensors and data driven modeling. Front Comput Neurosci 2018 Sep 11;12:72 [FREE Full text] [doi: 10.3389/fncom.2018.00072] [Medline: 30254580]
46. Schukat M, McCaldin D, Wang K, Schreier G, Lovell NH, Marschollek M, et al. Unintended consequences of wearable sensor use in healthcare. Contribution of the IMIA wearable sensors in healthcare WG. Yearb Med Inform 2016 Nov 10(1):73-86 [FREE Full text] [doi: 10.15265/IY-2016-025] [Medline: 27830234]
47. Folador JP, Vieira MF, Pereira AA, Andrade AO. Open-source data management system for Parkinson's disease follow-up. PeerJ Comput Sci 2021 Feb 17;7:e396 [FREE Full text] [doi: 10.7717/peerj-cs.396] [Medline: 33817042]
48. Hartson R, Pyla PS. The UX Book: Process and Guidelines for Ensuring a Quality User Experience. Amsterdam, The Netherlands: Elsevier Science; 2012.
49. Taktak A, Ganney P, Long D, Axell R. Clinical Engineering: A Handbook for Clinical and Biomedical Engineers. Amsterdam, The Netherlands: Elsevier Science; 2019.
50. Garzo A, Silva PA, Garay-Vitoria N, Hernandez E, Cullen S, Cochen De Cock V, et al. Design and development of a gait training system for Parkinson's disease. PLoS One 2018 Nov 12;13(11):e0207136 [FREE Full text] [doi: 10.1371/journal.pone.0207136] [Medline: 30418993]
51. Maramba I, Chatterjee A, Newman C. Methods of usability testing in the development of eHealth applications: a scoping review. Int J Med Inform 2019 Jun;126:95-104. [doi: 10.1016/j.ijmedinf.2019.03.018] [Medline: 31029270]
52. Phillippi J, Lauderdale J. A guide to field notes for qualitative research: context and conversation. Qual Health Res 2018 Feb;28(3):381-388. [doi: 10.1177/1049732317697102] [Medline: 29298584]
53. Braun V, Clarke V. Using thematic analysis in psychology. Qual Res Psychol 2006;3(2):77-101. [doi: 10.1191/1478088706qp063oa]
54. Lewis JR, Sauro J. Item benchmarks for the system usability scale. J Usability Stud 2018 May;13(3):158-167 [FREE Full text]
55. Bangor A, Kortum PT, Miller JT. An empirical evaluation of the system usability scale. Int J Hum Comput Interact 2008;24(6):574-594. [doi: 10.1080/10447310802205776]
56. Cosford K, Briere J, Ambros B, Beazley S, Cartwright C. Effect of instructional format on veterinary students' task performance and emotional state during a simulation-based canine endotracheal intubation laboratory: handout versus video. J Vet Med Educ 2020 Apr;47(2):239-247. [doi: 10.3138/jvme.0618-077r1] [Medline: 31194627]
57. Shah RF, Gupta RM. Video instruction is more effective than written instruction in improving inhaler technique. Pulm Pharmacol Ther 2017 Oct;46:16-19. [doi: 10.1016/j.pupt.2017.08.005] [Medline: 28797611]
58. Mays KK, Katz JE, Groshek J. Mediated communication and customer service experiences. Psychological and demographic predictors of user evaluations in the United States. Period Polytech Soc Manag Sci 2022 Jan 03;30(1):1-11. [doi: 10.3311/PPso.16882]
59. Kargar N, Choobineh AR, Razeghi M, Keshavarzi S, Meftahi N. Posture and discomfort assessment in computer users while using touch screen device as compared with mouse-keyboard and touch pad-keyboard. Work 2018;59(3):341-349. [doi: 10.3233/WOR-182685] [Medline: 29630576]
60. Lackermair G, Kailer D, Kanmaz K. Importance of online product reviews from a consumer's perspective. Adv Econ Bus 2013;1(1):1-5. [doi: 10.13189/aeb.2013.010101]
61. Binyamin S, Rutter M, Smith S. The utilization of system usability scale in learning management systems: a case study of Jeddah community college. In: Proceedings of the 9th annual International Conference of Education, Research and Innovation. 2016 Presented at: ICERI2016; November 14-16, 2016; Seville, Spain. [doi: 10.21125/iceri.2016.2290]
62. Nielsen J. Why you only need to test with 5 users. Nielsen Norman Group. 2000 Mar 18. URL: https://www.nngroup.com/articles/why-you-only-need-to-test-with-5-users/ [accessed 2022-08-22]
63. Virzi RA. Refining the test phase of usability evaluation: how many subjects is enough? Hum Factors 1992;34(4):457-468. [doi: 10.1177/001872089203400407]
64. Moradian S, Krzyzanowska MK, Maguire R, Morita PP, Kukreti V, Avery J, et al. Usability evaluation of a mobile phone-based system for remote monitoring and management of chemotherapy-related side effects in cancer patients: mixed-methods study. JMIR Cancer 2018 Dec 21;4(2):e10932 [FREE Full text] [doi: 10.2196/10932] [Medline: 30578238]
65. Fetters MD, Curry LA, Creswell JW. Achieving integration in mixed methods designs-principles and practices. Health Serv Res 2013 Dec;48(6 Pt 2):2134-2156 [FREE Full text] [doi: 10.1111/1475-6773.12117] [Medline: 24279835]
66. Green CA, Duan N, Gibbons RD, Hoagwood KE, Palinkas LA, Wisdom JP. Approaches to mixed methods dissemination and implementation research: methods, strengths, caveats, and opportunities. Adm Policy Ment Health 2015 Sep;42(5):508-523 [FREE Full text] [doi: 10.1007/s10488-014-0552-6] [Medline: 24722814]
67. Al-Tahat KS. Arabic translation, cultural adaptation and psychometric validation of the post-study system usability questionnaire (PSSUQ). Int J Hum Comput Interact 2021 May 11;37(19):1815-1822. [doi: 10.1080/10447318.2021.1913844]
68. Cornet VP, Toscos T, Bolchini D, Rohani Ghahari R, Ahmed R, Daley C, et al. Untold stories in user-centered design of mobile health: practical challenges and strategies learned from the design and evaluation of an app for older adults with heart failure. JMIR Mhealth Uhealth 2020 Jul 21;8(7):e17703 [FREE Full text] [doi: 10.2196/17703] [Medline: 32706745]

XSL•FO
RenderX

69. Graham AK, Munson SA, Reddy M, Neubert SW, Green EA, Chang A, et al. Integrating user-centered design and behavioral science to design a mobile intervention for obesity and binge eating: mixed methods analysis. JMIR Form Res 2021 May 10;5(5):e23809 [FREE Full text] [doi: 10.2196/23809] [Medline: 33970114]
70. Moldestad M, Petrova VV, Tirtanadi K, Mishra SR, Rajan S, Sayre G, et al. Improving the usability of written exposure therapy for therapists in the department of veterans affairs telemental health: formative study using qualitative and user-centered design methods. JMIR Form Res 2023 Nov 06;7:e47189 [FREE Full text] [doi: 10.2196/47189] [Medline: 37930747]
71. Sonney J, Cho EE, Zheng Q, Kientz JA. Refinement of a parent-child shared asthma management mobile health app: human-centered design study. JMIR Pediatr Parent 2022 Feb 17;5(1):e34117 [FREE Full text] [doi: 10.2196/34117] [Medline: 35175214]
72. Salinas E, Cueva R, Paz F. A systematic review of user-centered design techniques. In: Proceedings of the Design, User Experience, and Usability. Interaction Design. 2020 Presented at: DUXU 2020; July 19-24, 2020; Copenhagen, Denmark. [doi: 10.1007/978-3-030-49713-2_18]
73. John W, Bridges B, Taylor J. Evaluation of the Parkinson's remote interactive monitoring system (PRIMS): a usability study. JMIR Preprint posted online October 31, 2023 2023 [FREE Full text] [doi: 10.2196/54145]

Abbreviations

MDS-UPDRS: Movement Disorder Society–sponsored revision of the Unified Parkinson's Disease Rating Scale
PD: Parkinson disease
PRIMS: Parkinson's Remote Interactive Monitoring System
RA: research assistant
SUS: System Usability Scale

Edited by P Santana-Mancilla; submitted 31.10.23; peer-reviewed by I Straand, J Edwards; comments to author 22.01.24; revised version received 15.03.24; accepted 14.04.24; published 24.05.24.

Please cite as:
Bridges B, Taylor J, Weber JT
Evaluation of the Parkinson's Remote Interactive Monitoring System in a Clinical Setting: Usability Study
JMIR Hum Factors 2024;11:e54145
URL: https://humanfactors.jmir.org/2024/1/e54145
doi:10.2196/54145
PMID:38787603

©Bronwyn Bridges, Jake Taylor, John Thomas Weber. Originally published in JMIR Human Factors (https://humanfactors.jmir.org), 24.05.2024. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on https://humanfactors.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Assessing the Relationship Between Digital Trail Making Test Performance and IT Task Performance: Empirical Study

Tanguy Depauw[1], MSc; Jared Boasen[1,2], PhD; Pierre-Majorique Léger[1], PhD; Sylvain Sénécal[1], PhD

[1]Tech3lab, HEC Montréal, Montréal, QC, Canada

[2]Faculty of Health Sciences, Hokkaido University, Sapporo, Hokkaido, Japan

Corresponding Author:
Jared Boasen, PhD
Tech3lab
HEC Montréal
5540 Av Louis-Colin
Montréal, QC, H3T 1T7
Canada
Phone: 1 514 340 6000
Email: jared.boasen@hec.ca

Abstract

Background: Cognitive functional ability affects the accessibility of IT and is thus something that should be controlled for in user experience (UX) research. However, many cognitive function assessment batteries are long and complex, making them impractical for use in conventional experimental time frames. Therefore, there is a need for a short and reliable cognitive assessment that has discriminant validity for cognitive functions needed for general IT tasks. One potential candidate is the Trail Making Test (TMT).

Objective: This study investigated the usefulness of a digital TMT as a cognitive profiling tool in IT-related UX research by assessing its predictive validity on general IT task performance and exploring its discriminant validity according to discrete cognitive functions required to perform the IT task.

Methods: A digital TMT (parts A and B) named Axon was administered to 27 healthy participants, followed by administration of 5 IT tasks in the form of CAPTCHAs (Completely Automated Public Turing tests to Tell Computers and Humans Apart). The discrete cognitive functions required to perform each CAPTCHA were rated by trained evaluators. To further explain and cross-validate our results, the original TMT and 2 psychological assessments of visuomotor and short-term memory function were administered.

Results: Axon A and B were administrable in less than 5 minutes, and overall performance was significantly predictive of general IT task performance ($F_{5,19}$=6.352; P=.001; Λ=0.374). This result was driven by performance on Axon B ($F_{5,19}$=3.382; P=.02; Λ=0.529), particularly for IT tasks involving the combination of executive processing with visual object and pattern recognition. Furthermore, Axon was cross-validated with the original TMT (P_{corr}=.001 and P_{corr}=.017 for A and B, respectively) and visuomotor and short-term memory tasks.

Conclusions: The results demonstrate that variance in IT task performance among an age-homogenous neurotypical population can be related to intersubject variance in cognitive function as assessed by Axon. Although Axon's predictive validity seemed stronger for tasks involving the combination of executive function with visual object and pattern recognition, these cognitive functions are arguably relevant to the majority of IT interfaces. Considering its short administration time and remote implementability, the Axon digital TMT demonstrates the potential to be a useful cognitive profiling tool for IT-based UX research.

(JMIR Hum Factors 2024;11:e49992) doi:10.2196/49992

KEYWORDS

Trail Making Test; user experience; cognitive profile; information technology; task performance; cognitive assessment; human factors; cognitive function; CAPTCHA

XSL•FO
RenderX

Introduction

Cognitive functional ability is a fundamental factor widely recognized to influence IT usability [1-3]. The classical approach to control for cognitive functional ability is to target participants according to general demographics based on age, education, or other factors [4,5]. However, this approach intrinsically precludes the ability to control for or assess how cognitive functional ability impacts IT usability in individual users, thereby limiting the extent which insight can be gained within a demographic or for an individual. Moreover, this approach is incongruent with the rapid advancement of IT toward products that adapt to individual user characteristics, thus necessitating a more granular understanding of individual cognitive abilities [6-8].

To obtain a granular characterization of individual cognitive function, hitherto, research has typically used cognitive assessment batteries [9-11]. Dumont et al [12] used the National Institutes of Health Toolbox, which is a battery of cognitive tests that can be completed in 40 minutes [13] to develop a cognitive analysis grid to be able to draw statistical parallels between the cognitive demands of an information systems interface and the performance of a user. Other batteries of tests were also used, such as the Kit of Factor-Referenced Cognitive Tests [10], which was used by Wagner et al [1] to study the impact of age on website usability and by Allen [14] in his research to study the combination of users' cognitive abilities and specific information system functionalities that can be implemented to create system usability. This battery is typically administered in 144 minutes [15]. Another approach for assessing individual cognitive ability is to use clinically administered tests such as the Montreal Cognitive Assessment (MoCA) or the Mini-Mental State Examination (MMSE). Although typically used in medical settings to evaluate cognitive impairment in patients with neurological disorders [9,11], MoCA and MMSE have been reportedly used to measure the cognitive abilities of participants in human-computer interaction experiments [3,16-18]. However, while detailed and accurate, these cognitive assessment batteries are too lengthy to practically administer during typical user experience (UX) testing time frames [19,20]. Furthermore, while clinically administered tests such as MoCA and MMSE are comparatively shorter than other assessment batteries, they require a trained administrator to administer and score the test [3]. This level of expertise may not always be available, particularly in UX research settings where mostly nonclinically trained research personnel are conducting the experiments.

Correspondingly, there have been calls from across health, UX, and IT domains for a more practical yet accurate means of assessing cognitive function [12,21,22]. One solution would be to identify a short test with reduced scope but which nevertheless targets cognitive functions important for using IT. Based on research conducted to understand the impact of cognitive functions on the use of technology by older people [23,24], and on existing models of cognitive architecture in human-computer interaction [25], we identified 5 key cognitive functions important for IT use: visual perception, motor function, executive function, inhibitory control, and working memory. Visual perception is important for finding relevant information cues on a web page [23]. Motor functions are involved in tasks such as data entry using the keyboard, navigation using the mouse, or other tool to perform a digital task [26]. Executive functions come into play in order to make decisions and prioritize action [23]. Inhibitory control, also called "response inhibition" [27], is the functional ability to inhibit or override motor commands or other executive processing, such as when an external stimulus interferes with goal-driven behavior as in a task-switching situation [28,29]. Finally, short-term or working memory capacity may be important in IT task performance, for example, for remembering options or system output at a later stage [23].

One potential preexisting cognitive assessment candidate that targets these cognitive functions related to IT use is the Trail Making Test (TMT). First developed for the Army Individual Test Battery [30], the TMT is one of the most widely used instruments in neuropsychological assessment as an indicator of cognitive processing speed and executive functioning [31-35]. Many studies have been conducted to determine which cognitive abilities are engaged during the completion of this 2-part test (TMT-A and TMT-B). After a comprehensive review of the literature on the topic, Sánchez-Cubillo et al [36] explored the contributions of certain cognitive functions and found that part A of the TMT (TMT-A) mainly requires visual-perceptual abilities, and that part B (TMT-B) reflects primarily working memory, executive function, and task-switching ability. Finally, although its contribution in the TMT has been questioned by the study of Sánchez-Cubillo et al [36], it is interesting to note that psychomotor ability has been mentioned numerous times as one of the abilities required for both parts of the TMT (Groff and Hubble [37] in both parts, Schear and Sato [38], Gaudino et al [39], and Crowe [40] in part B). The primary objective of this study was to test the validity of using the TMT as a cognitive profiling tool to predict or explain the variance in IT task performance. With an interest in a practical tool for cognitive profile assessments in UX testing of digital artifacts, we chose to use a digital version of the TMT. To further support and explain our results, we additionally cross-validated the digital TMT with the original TMT, a visual search task assessing visuomotor processing [41,42], and a hidden path learning task assessing visuomotor-processing speed, spatial working memory, and error-monitoring ability [43]. We had two hypotheses: (1) TMT times would be predictive of general IT task performance and (2) that the predictive power of the TMT would be stronger for tasks requiring the use of cognitive functions that are congruent with those assessed by the TMT.

Methods

Sample

To test our hypothesis, we conducted a laboratory experiment with 27 healthy participants (12 men and 15 women), between 18 and 36 (mean 24, SD 4.22) years of age, who were mostly university students (n=22, 85%).

Ethical Considerations

Written informed consent was obtained from all subjects via a signed form at the beginning of the experiment. This project

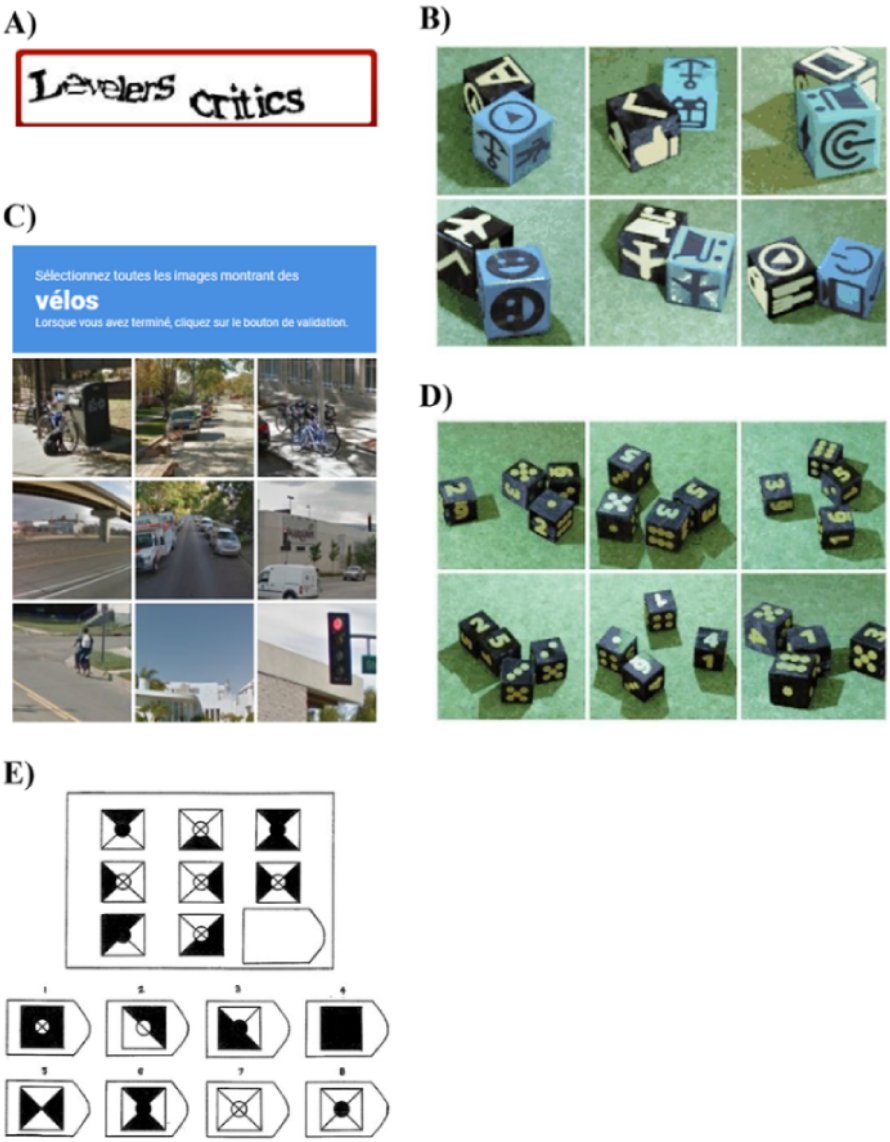
XSL•FO
RenderX

was approved by our institution's research ethics committee (#2021-4108). A monetary compensation of CAD $25 (US $18.35) was provided to each subject upon completion of the experiment. Data from 1 subject were lost due to technical issues, thus leaving data from 26 participants available for analysis. All data were anonymized prior to analysis and stored in encrypted servers only accessible by authorized researchers.

IT Tasks

Two types of general IT tasks were used in the experiment. One type of IT task was based on CAPTCHA (Completely Automated Public Turing Tests to Tell Computers and Humans Apart). This type of Turing test is widely used in IT to ensure the cybersecurity of many internet services, as they prevent a number of attacks from automated programs (often referred to as bots), by distinguishing legitimate users from computer bots while requiring minimal effort by the human user [44]. Four CAPTCHAs were based on typical existent CAPTCHAs and included Google reCAPTCHA (Google), pictogram recognition (PicRec), numerical recognition (NumRec), and text recognition (Text). A Fifth task was taken from Raven's Progressive Matrices (RPM) and presented in a CAPTCHA format. RPM are a collection of widely used standardized intelligence tests consisting of analogy problems in which a matrix of geometric figures is presented with 1 entry missing, and the correct missing entry must be selected from a set of answer choices [45]. A 3×3 RPM was selected as it was considered that it offered the best trade-off between cognitive effort and the time required to complete it. The final 5 IT tasks, shown in Figure 1, were embedded on a Qualtrics questionnaire. For this study, we targeted IT task completion time, measured as the time from the display of each task to when subjects responded and pressed the "next" button, based on 30 fps screen recordings.

Figure 1. The 5 information technology tasks. (A) Text-based Completely Automated Public Turing tests to tell Computers and Humans Apart (CAPTCHA): subjects had to type the 2 words in an input field below the text image. (B) Pictogram recognition CAPTCHA: subjects had to recognize and click on the image showing the 2 dice with the same pictogram on the top face. (C) Google reCAPTCHA: subjects had to recognize and click on the images showing the bicycles. (D) Number recognition CAPTCHA: subjects had to recognize and click on the image showing dice summing to 14 on the top faces (numerals and dots combined). (E) Raven's Progressive Matrix: subjects had to click from among the 8 proposed images the one which most appropriately fit in the missing corner of the basic matrix.



XSL•FO
RenderX

The other type of IT task was a website design evaluation to assess perceived usability using Aladwani and Palvia's [46] user-perceived web quality measurement scale. Screenshots of the home pages of the following 5 websites were used: Vignerons d'Exception [47], Renaud-Bray [48], LesPAC [49], [50], and [51]. One website was presented subsequent to each CAPTCHA. Participants were told that the website evaluation was the primary task of the experiment and that the CAPTCHAs were present as a security measure to access our database housing the website screenshots. However, the website evaluations were actually dummy tasks, and participant responses were not analyzed. The IT tasks really targeted and analyzed in this study were the CAPTCHAs.

Cognitive Function Characterization of CAPTCHAs

The principal reason CAPTCHAs were chosen as our general IT tasks is because they are ubiquitous in IT and because they are often distinguishable from one another according to task-specific demands such as math, 3D orientation, text recognition, and visual search, suggesting that different underlying cognitive processing required them. However, there is a paucity of studies regarding the examination of the specific cognitive functions of CAPTCHAs. Therefore, we formed a panel of 11 trained, nonexpert evaluators to rank the selected CAPTCHAs on a 5-point agreement scale according to the 5 cognitive functions mentioned in the Introduction section, which have been deemed relevant to IT tasks and the TMT: visuospatial perception, motor function, executive function, inhibitory control, and working memory. The evaluation scores permitted each CAPTCHA to be assigned a rank according to the extent the cognitive functions required to perform it overlapped with those of the TMT. In order of highest to lowest alignment, the rankings were as follows: (1) RPM, (2) NumRec, (3) PicRec, (4) Google, and (5) Text, as shown in Table 1. For details of how this evaluation was conducted and how the process was validated, see Multimedia Appendix 1.

Table 1. Convergence ranks of IT tasks with the TMT[a].

IT task	RPM[b] (E)	NumRec[c] (D)	PicRec[d] (B)	Google (C)	Text (A)
Executive function, mean (SD)	5.00 (0.00)	4.91 (0.30)	4.45 (1.04)	4.45 (1.04)	3.82 (1.4)
Visual object recognition, mean (SD)	4.09 (0.70)	4.27 (0.65)	4.64 (0.67)	4.82 (0.60)	4.18 (1.17)
Visual pattern recognition, mean (SD)	4.91 (0.30)	4.45 (0.93)	4.64 (0.67)	3.82 (0.98)	4.64 (0.67)
Working memory, mean (SD)	4.18 (0.60)	3.91 (1.38)	2.91 (1.51)	2.45 (1.21)	2.73 (1.27)
Evaluation score for reliable convergent dimensions, mean (SD)	4.55 (0.48)	4.39 (0.42)	4.16 (0.84)	3.89 (1.04)	3.84 (0.81)
Convergence rank with TMT following the evaluation[e]	1	2	3	4	5

[a]TMT: Trail Making Test.

[b]RPM: Raven's Progressive Matrices.

[c]NumRec: numerical recognition.

[d]PicRec: pictogram recognition.

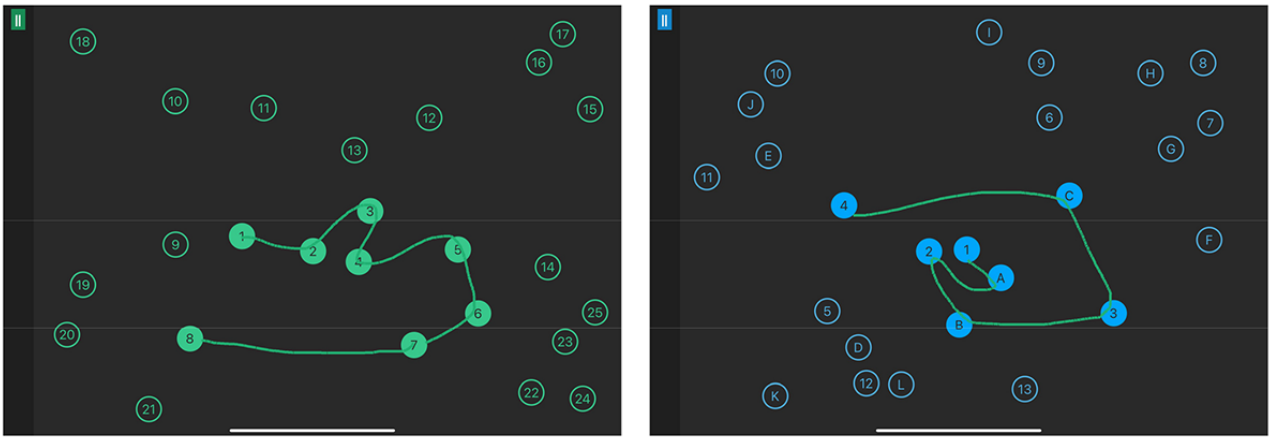
[e]Based on the average evaluation scores of IT tasks on the reliable cognitive dimensions considered convergent with the TMT. A, B, C, D, and E refer to the labels of the IT tasks presented in Figure 1.

Digital TMT

Because we are interested in cognitive assessment for UX testing of IT and because it was convenient to present all the tasks on the same device, we chose to use a digital version of the TMT called "Axon" (Language Research Development Group). This version emulates the original TMT as an iPad app, allowing the user to draw the trail on the touch screen with 1 finger. The 2 parts (A and B) of the TMT were completed, each with 25 circles to connect. Axon TMT was designed with a canvas generation algorithm, meaning that the test canvas for each subject for each TMT A and B was different. As shown in Figure 2, both tests were presented in full screen on the iPad with 25 circles of 1-cm diameter placed randomly on the digital canvas in a homogeneous way. The rules of Axon were identical to those of the original TMT, as outlined by Bowie and Harvey [52]. Participants had to connect the circles in ascending order: from 1 to 25 for part A and from 1 to 13 for part B, alternating numbers and letters in ascending order (ie, 1, A, 2, B, 3, C, etc). Errors such as lifting the finger off the screen, crossing trails, or connecting a wrong circle resulted in the line for the latest segment to be automatically erased and subjects had to return to the last successfully reached circle in order to continue. The measures chosen for this study were the completion time for each of the 2 parts of the test, from the moment the layout was displayed until the last circle was reached. These measures were exported from the app after the completion of the study and used in our statistical analyses.

XSL•FO
RenderX

Figure 2. Screenshots of Axon A and Axon B. Subjects had to draw to connect the circles in ascending order (from 1 to 25 for part A and from 1 to 13 and A to L for part B, alternating numbers and letters) on a single line, without crossing paths or lifting their finger from the screen. In case of errors in drawing, the app automatically guided subjects back to the last correct circle from which they continued the test.



Cross-Validation of the Digital TMT

Overview

To better support and explain our results, we cross-validated Axon with the original TMT and a working memory and a visual search task.

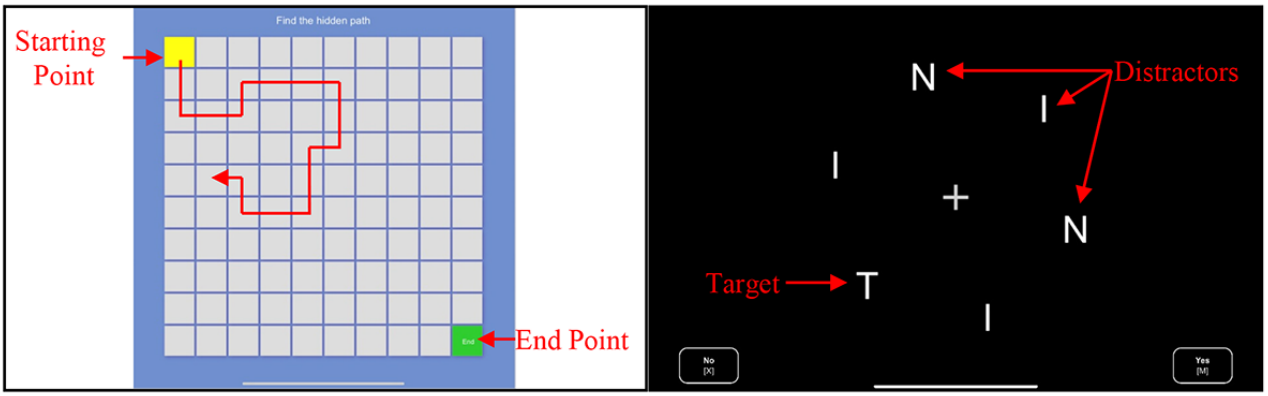
Original TMT

The original TMT was administered as outlined by Bowie and Harvey [52] at the end of the study. The practice step was skipped in the interest of time and with the knowledge that the subject had already performed the digital TMT earlier in the study.

Hidden Path Learning Task

To cross-validate Axon's ability to measure working memory and spatial ability, we administered a hidden maze learning task, based on the Groton Maze Learning Test developed by Pietrzak et al [43]. Our task was called the "hidden path learning" task and was based on a 10 × 10 grid. Five trials were administered on the iPad via the Cognition Lab platform (BeriSoft, Inc), following similar guidelines as the Groton Maze Learning Test [43]. The hidden path learning task is particularly targeted at working memory, as the user has to call on it to navigate between tiles and remember any errors they may have made before [53,54]. Correspondingly, working memory ability is associated with the extent to which completion time decreases over trials, revealing a learning curve. Thus, the metrics used for these analyses were the difference between the completion times of each consecutive trial on the task. A depiction of the hidden path learning task is shown in Figure 3 (left). Measures were automatically collected on the Cognition Lab server.

Figure 3. Cross-validation tasks. In the hidden path learning 10×10 matrix (left), subjects had to go from the yellow starting point to the green end point 1 tile at a time. In the visual search task (right), there were 6 items, with 1 target and 5 distractors. In the I+N sequence (shown), participants had to touch "Yes" at the bottom-right if they saw the target, "No" at the bottom-left otherwise.



Visual Search Task

To cross-validate Axon's ability to measure visuomotor function, we administered a visual search task on the Cognition Lab platform (BeriSoft, Inc). This task was based on the work by Treisman and Gelade [42] and involved finding a target among distractors. Participants had to touch the right side of the screen when they saw the target, the left side otherwise, therefore involving visual and psychomotor response ability. Three stimuli configurations were used, with 3 distractor sets. Configurations were displayed with 24 trials for each stimulus, leading to a total of 72 trials. For each trial, 3, 6, or 9 symbols were displayed (letters or shapes), with even and randomized distribution among each stimulus sequence. A depiction of this task is shown in Figure 3 (right). Again, measures were

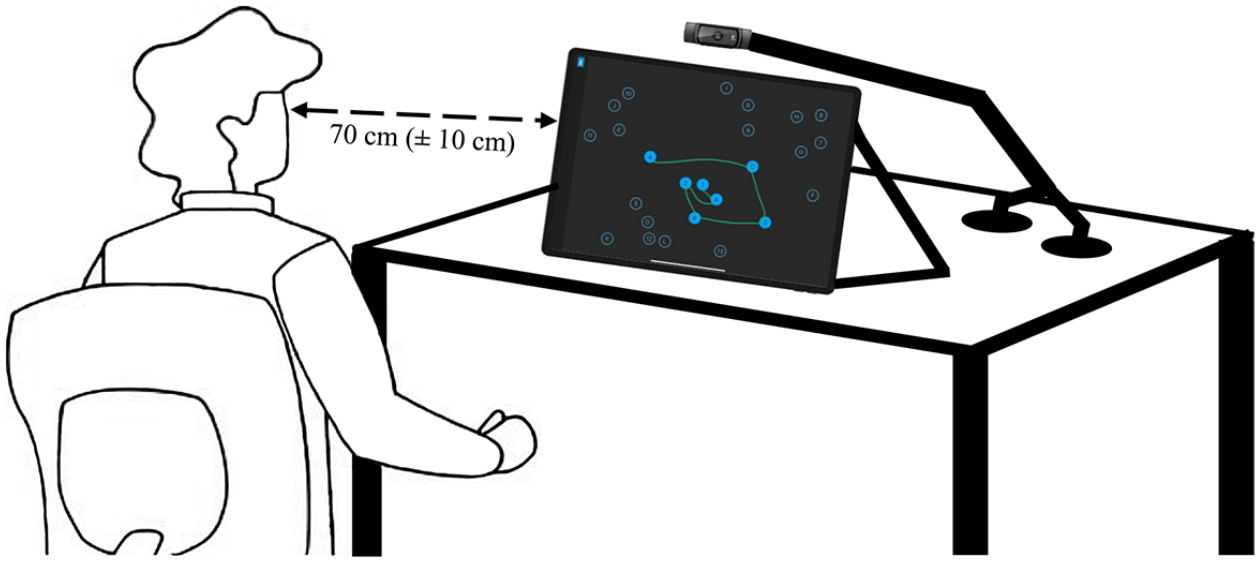
XSL•FO
RenderX

automatically collected on the Cognition Lab server. Reaction times were used for the present analyses.

Procedures

Upon arrival and after signing the informed consent form, subjects were asked to sit on a chair facing the iPad Air (fourth generation) running on iPadOS 15.3 (Apple Inc) placed on a desk and were asked to adjust the chair's height so that they were comfortable using the iPad, and they were within the camera recording frame. The experimental setup is presented in Figure 4. They were asked to move the chair closer or further away to maintain an approximate distance of 70 (±10) cm between their eyes and the iPad screen to give enough space for hand movement during the tasks. The camera was fixed independently from the iPad to avoid unwanted movements on the video when the participant presses the screen while doing the tasks. After a presentation of the study and the tools used, the participants were asked to complete the 2 parts of the TMT (A and then B) on the Axon app. Task instructions were given in a protocol format to ensure that all participants received the same instructions and that the data would be comparable. Participants were verbally and visually guided through the rules of the TMT using a tutorial embedded in the app.

Figure 4. Experimental setup diagram. The subject was seated at a chair in front of a desk where the iPad Air 4 was placed. A Logitech C920 camera was independently fixed to the desk via a camera stand and duct tape.



After completing parts A and B on the Axon app, participants were administered the hidden path learning and the visual search tasks. Then, participants commenced the IT task portion of the experiment. As previously mentioned, participants were told that the primary objective was to evaluate 5 interfaces of more or less popular websites, each interface being on a secure server accessible only after the completion of a CAPTCHA. Thus, subjects completed a CAPTCHA, observed a web interface for a few minutes, and then completed the user-perceived web quality measurement scale [46]. This sequence was repeated 5 times, with the tasks presented in random order, each preceded by a distinct CAPTCHA. At the end of the study, for ethical reasons, subjects were told orally that they were in fact being evaluated on their performance on the CAPTCHAs.

Statistical Analyses

To test the ability of the Axon TMT to predict performance on the 5 CAPTCHA IT tasks, a repeated-measures multivariate analysis of covariance (RM MANCOVA) was performed with Axon A completion time and Axon B completion time as independent predictors and the completion time for each of the 5 IT tasks as the dependent covariates.

To further interpret our results, we tested the relationship between Axon TMT completion times and visuomotor function by performing an RM MANCOVA with Axon A and Axon B times as independent predictors and the mean reaction time of each of the 3 visual search tasks (the shape of an arrow as a target among the triangle shapes as distractors, the letter T as a target among the letters I and N as distractors, and the letter T as a target among the letters I and Z as distractors) as the dependent covariates. In addition, we tested the relationship between Axon TMT completion times and working memory function by performing an RM MANCOVA with Axon A and Axon B times as independent predictors and the difference between the completion time of each consecutive trial on the hidden path learning task as the dependent covariates. Finally, we cross-validated the relationship between the Axon TMT and the original TMT using 2 Pearson correlation tests, 1 each for tests A and B.

For all RM MANCOVAs performed in the analysis, omnibus results and multivariate results for each independent predictor are reported. In the case of significant multivariate results, simple main effects based on parameter estimates are reported for dependent covariates, which were significantly predicted by Axon.

All statistical analyses were conducted using the IBM SPSS Statistics software (version 28.0.1.1; IBM Corp) with a threshold for statistical significance set at $P \leq .05$, using the Bonferroni correction to adjust for multiple comparisons.

XSL•FO
RenderX

Results

Axon TMT Cross-Validation

Axon Versus Original TMT

The mean scores of Axon A and B were 48.04 (SD 25.80) and 56.88 (SD 25.53) seconds, respectively. The mean scores on the original TMT A and B were 29.22 (SD 12.26) and 51.62 (SD 19.07) seconds, respectively. Pearson correlation tests revealed that Axon is highly correlated with TMT results, with a significant positive correlation between Axon A and TMT A (r=0.688; P_{corr}=.001) and a significant positive correlation between Axon B and TMT B (r=0.505; P_{corr}=.017).

Axon TMT Versus Hidden Path Learning

The difference in consecutive trial times was (2–1) –29.87 (17.70), (3–2) –5.48 (6.01) seconds, (43) –4.30 (4.80) seconds, and (5–4) –1.50 (4.25) seconds. The omnibus test of the RM MANCOVA revealed that Axon A and Axon B combined are significant to explain the variance in the decrease in completion times across consecutive trials ($F_{4,20}$=4.119; P=.01; Λ=0.548). However, multivariate results revealed that the decrease in completion times across trials was not predicted by Axon A ($F_{4,20}$=1.923; P=.15; Λ=0.722) or Axon B ($F_{4,20}$=1.106; P=.38; Λ=0.819) alone. Thus, a predictive relationship appears to exist between Axon and working memory in the hidden path learning task as a function of Axon A and B combined.

Axon TMT Versus Visual Search

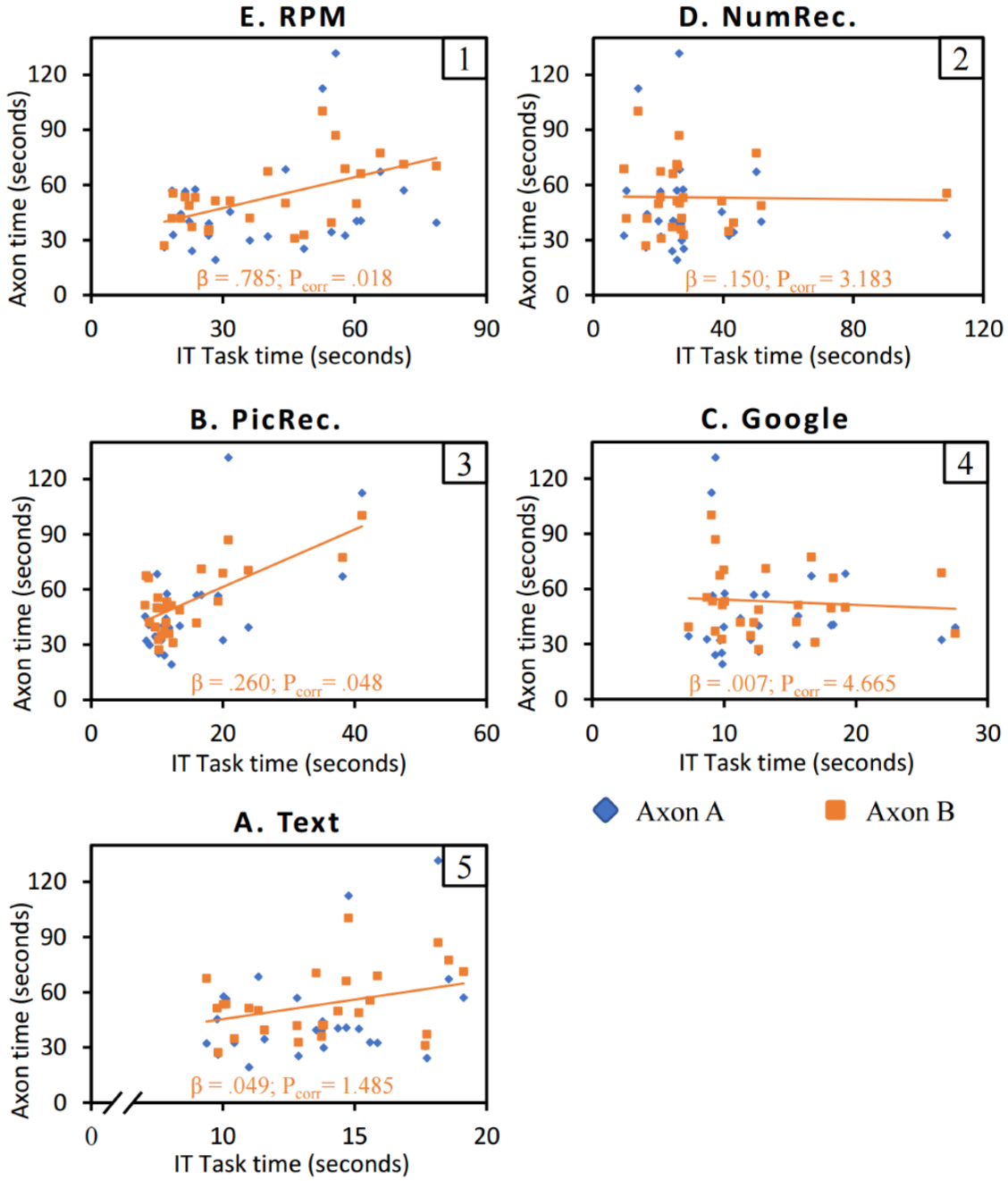
Reaction times for the T among letters I and N, T among letters I and Z, and arrow among triangles were 0.80 (0.14) milliseconds, 0.78 (0.15) milliseconds, and 0.68 (0.14) milliseconds, respectively. The omnibus test of the RM MANCOVA revealed that Axon A and Axon B combined significantly explained the variance in visuomotor function assessed with reaction time to the 3 stimuli in the visual search task ($F_{3,21}$=3.125; P=.048; Λ=0.691). Multivariate results revealed that this result was driven mainly by Axon A ($F_{3,21}$=3.220; P=.043; Λ=0.685) rather than Axon B ($F_{3,21}$=0.502; P=.69; Λ=0.933). Parameter estimates revealed that Axon A was marginally significantly predictive of reaction times to the letter T among letters I and N stimulus (β=3.573; t_{21}=2.767; P_{corr}=.055) and significant for letter T among letters I and Z (β=4.353; t_{21}=3.156; P_{corr}=.02) and arrow among triangles (β=3.725; t_{21}=3.158; P_{corr}=.02) stimuli.

Axon TMT Predicts Overall IT Performance

The primary hypothesis assumed that there was a positive predictive relationship between TMT performance and IT task performance. The omnibus test of the RM MANCOVA revealed that Axon A and Axon B combined significantly explain the variance in IT tasks performance ($F_{5,19}$=6.352; P=.001; Λ=0.374), thereby supporting the primary hypothesis. Multivariate results revealed that this effect was driven by performance on Axon B ($F_{5,19}$=3.382; P=.03; Λ=0.53). Figure 5 shows the distribution of Axon completion times in relation to IT task completion times.

XSL•FO
RenderX

Figure 5. Distribution of Axon A and B completion times in relation with the completion times of the 5 IT tasks (N=26). Axon B trendlines and parameter estimates (β and P) show the relationship between Axon B IT task performance. Number in upper right corner of plot area is hypothesized convergence rank (Table 1). IT: information technology; NumRec: numerical recognition; PicRec: pictogram recognition; RPM: Raven's Progressive Matrices. Letters A through E refer to the labels used for each task in Figure 1.



Axon TMT Better Predicts Performance on Convergent IT Tasks

The second hypothesis assumed that the predictive relationship between TMT performance and IT task performance would be stronger if the cognitive abilities involved in the performance were congruent. To test our hypothesis, we analyzed the parameter estimates for the multivariate results of Axon B. These revealed that Axon B was significantly predictive of IT task C (RPM task; β=.785; t_{19}=3.240; P_{corr}=.018) and IT task B (PicRec task; β=.260; t_{19}=2.824; P_{corr}=.048). However, IT task D (number recognition task), which was rated the second most congruent task with Axon, was not significantly predicted by Axon B (β=.150; t_{19}=0.479; P_{corr}=3.183). Our secondary hypothesis is therefore partially supported. These results are shown in Figure 5, where the effects of individual factors of Axon B on performance on IT tasks are represented (β and *P* values).

Discussion

Principal Findings

Cognitive functional ability may well affect task performance in UX and other research experimentation, leading to variance in performance measures among the target population and confounding the effects of experimental factors. Although

XSL•FO
RenderX

detailed cognitive assessment batteries exist and can be used to control intersubject differences in cognitive abilities [12], they are not time efficient and thus impractical to implement within typical experimental time frames. Here, this study tested the validity of using the Axon TMT, which takes only a few minutes to administer, to predict or explain the variance in IT task performance in an age-homogenous subject population.

The mean age of the subject population of this sample was 24 (SD 4.22) years. This is typical of many research studies, UX related or otherwise, relying on student recruitment through the parent institution [55-57]. Despite the relatively low SD of age, the SD in Axon TMT scores was broad, at 25.80 (mean 48.04) and 25.53 (mean 56.9) seconds, respectively, for Axon A and B, suggesting a large distribution of cognitive functional abilities among this age-homogeneous neurotypical population. Notably, the means and SDs for the Axon TMT, particularly for Axon A, were higher than what is typically reported in the literature for neurotypical subjects in this age bracket [58-60]. This may be due to the fact that, unlike in the implementation of the paper-based TMT, subjects did not practice a mini version of the test before performing Axon A or B. Thus, some portion of the time taken to complete the test must be attributable to familiarization with task demands. This would also explain why the mean scores for Axon B, whose task demands are similar to Axon A in many respects, are closer to typically reported TMT B means. Nevertheless, for the purposes of this study, it is not absolute Axon TMT scores that are important. Rather, it is the relative distribution of the variance in Axon scores and their correlation to other metrics that is essential. To that end, both Axon A and B significantly correlated to their respective paper-based TMT counterparts showed a combined predictive validity toward working memory via the hidden path learning task. Furthermore, it was Axon A, not B, which was the predominant driver of the significant correlation with visual search performance. This is logical, as the visual search task does not involve working memory–related processing [42,61]. Instead, it requires an emphasis on target identification, cognitive control, and motor output, precisely the dominant cognitive functions involved in TMT A [36,39,40]. Thus, far from being problematic, implementing Axon A and B without a preliminary minitest for practice was time-efficient and yielded a reliable distribution of scores, which could be cross-correlated with expected cognitive functions.

This cross-validation lends credibility to our observation that Axon A and B combined were significantly predictive of IT task performance, supporting our primary hypothesis. Interestingly, for the IT tasks chosen, it was Axon B that appeared as the stronger driver of predictive validity, suggesting that it may be more powerful in capturing the executive decision-making involved in an ecologically valid IT task. Moreover, simple main effects tests revealed that Axon B significantly correlated with 2 out of the top 3 tasks ranked as requiring congruent cognitive functions as the TMT, thereby partially supporting our secondary hypothesis. Contrary to our expectations, the NumRec task, which had the second-highest congruence rank, was not significantly correlated with Axon B. We speculate that the confound here relates to the underlying mathematical operations involved in solving that CAPTCHA. Although raters classified this as executive decision-making, it certainly can be said that neither TMT A nor TMT B requires arithmetic. Therefore, there must be cognitive processes involved that are simply not recruited during the performance of the TMT, which our ranking system was not granularized enough to capture, hence explaining the lack of correlation between the NumRec task and Axon B. Meanwhile, Axon B was most strongly correlated with the RPM and PicRec task, suggesting that it is well suited for tasks involving visual pattern and object recognition in combination with higher-order executive processing to orient this visual information. These kinds of processing are arguably crucial for interface navigation, virtual reality, gaming, or using simulators, which are extremely common IT tasks investigated in UX research [62-64]. Thus, while Axon does appear to be better aligned with IT tasks involving convergent cognitive processing, such tasks may well comprise a major proportion of those studied in UX research.

Finally, there are a few points worth emphasizing. First, the complete administration of Axon took less than 5 minutes, far shorter than the strategy used by Dumont et al [12] or any other cognitive assessment that we are aware of. Second, considering Axon's ability to differentiate from among an age-homogeneous neurotypical population, it would likely perform even better among populations where a larger variance in cognitive function would be expected, such as in older adults, children, stroke survivors, or other individuals with atypical cognitive function. This is important because understanding how to design appropriate and accessible IT for these populations has become a topic of increasing concern in UX research [65-67]. Moreover, Axon is suitable for remotely moderated experimentation, a popular strategy since the COVID-19 pandemic [68] and one that mitigates subject recruitment challenges for all population types. Finally, the current advancement in technology, particularly in the field of artificial intelligence, is trending toward a more personalized and user-centric approach, adapting technology to individual user characteristics such as preferences and interests [8,69,70]. Part of this personalization could be to tailor technology according to the cognitive abilities of users. Axon could potentially facilitate this advancement, serving as a quick and reliable metric to train the artificial intelligence technology adaptation algorithm.

Limitations

There are some limitations that should be acknowledged with this study. First, because the Axon app is designed to produce TMT canvases according to an algorithm with every test instance, the Axon A and B canvas layouts were not constant across subjects. This means that some of the variance in Axon A and B times is intrinsically attributable to factors such as differences in the straight-line drawing path length of the test or the extent of visual interference between each drawing segment. On the other hand, the fact that Axon A and B were significantly cross-validated with the original TMT and the visual search and hidden path learning tasks in spite of canvas layout differences between participants suggests that the variance these differences cause is small and does not detract from the use of Axon as a cognitive profiling tool in UX testing. Second, this study tested the predictive validity of Axon on simple and discrete IT tasks. This was necessary as a proof of

XSL•FO
RenderX

concept for our hypotheses. However, readers should use caution when generalizing the present results. Further research is needed to investigate the extent to which Axon retains predictive validity for more complex IT tasks in different contexts and across various user demographics, including neuroatypical and cognitively impaired users.

Conclusions

This study tested the ability of the Axon digital TMT to predict performance on discrete IT tasks. The results indicate that variance in IT task performance among an age-homogenous neurotypical population can be related to intersubject variance in cognitive function as assessed by Axon. Although the findings suggest that Axon's predictive validity may be strongest for IT tasks involving the combination of decision-making with visual object and pattern recognition, these types of cognitive processing would arguably be relevant to the majority of IT interfaces. Considering its short administration time and remote implementability, the Axon digital TMT has the potential to be a useful cognitive profiling tool for IT-based UX research.

Acknowledgments

The authors would like to thank David Brieugne, Salima Tazi, Xavier Côté, and all the operational staff at the Tech3Lab for their operational assistance in the execution of this study. This study was made possible thanks to funding from NSERC (ALLRP 571020-21), Prompt (176_Sénécal-BNC, CAE, CN, SRC_2021.02), and Mitacs (IT18280).

Conflicts of Interest

None declared.

Multimedia Appendix 1
Evaluation of CAPTCHAs–congruent cognitive functions in the information technology tasks.
[DOCX File , 17 KB - humanfactors_v11i1e49992_app1.docx]

References

1. Wagner N, Hassanein K, Head M. The impact of age on website usability. Computers in Human Behavior 2014 Aug;37:270-282. [doi: 10.1016/j.chb.2014.05.003]
2. Karahoca D, Karahoca A, Güngör A. Assessing effectiveness of the cognitive abilities and individual differences on e-learning portal usability evaluation. New York, NY, United States: Association for Computing Machinery; 2008 Presented at: Proceedings of the 9th International Conference on Computer Systems and Technologies and Workshop for PhD Students in Computing; June 12-13, 2008; Gabrovo, Bulgaria. [doi: 10.1145/1500879.1500942]
3. Martins AI, Silva AG, Pais J, Cruz VT, Rocha NP. The impact of users' cognitive function on evaluator perceptions of usability. Sci Rep 2022 Aug 12;12(1):13753 [FREE Full text] [doi: 10.1038/s41598-022-17441-3] [Medline: 35962003]
4. Aykin NM, Aykin T. Individual differences in human-computer interaction. Comput Ind Eng 1991;20(3):373-379. [doi: 10.1016/0360-8352(91)90009-U]
5. Lewis JR. Usability testing. In: Salvendy G, editor. Handbook of Human Factors and Ergonomics, Fourth Edition. Hoboken, NJ: John Wiley & Sons; 2012:1267-1312.
6. Ji H, Yun Y, Lee S, Kim K, Lim H. An adaptable UI/UX considering user's cognitive and behavior information in distributed environment. Cluster Comput 2017;21(1):1045-1058. [doi: 10.1007/s10586-017-0999-9]
7. Andreessen LM, Gerjets P, Meurers D, Zander TO. Toward neuroadaptive support technologies for improving digital reading: a passive BCI-based assessment of mental workload imposed by text difficulty and presentation speed during reading. User Model User-Adap Inter 2020;31(1):75-104 [FREE Full text] [doi: 10.1007/s11257-020-09273-5]
8. Gajos KZ, Weld DS, Wobbrock JO. Automatically generating personalized user interfaces with Supple. AI 2010;174(12-13):910-950 [FREE Full text] [doi: 10.1016/j.artint.2010.05.005]
9. Folstein MF, Folstein SE, McHugh PR. "Mini-mental state": a practical method for grading the cognitive state of patients for the clinician. J Psychiatr Res 1975;12(3):189-198. [doi: 10.1016/0022-3956(75)90026-6] [Medline: 1202204]
10. Ekstrom RB, French JW, Harman HH, Dermen D. Manual for Kit of Factor-Referenced Cognitive Tests, 1976. Princeton, NJ: Education Testing Service; 1976.
11. Nasreddine ZS, Phillips NA, Bédirian V, Charbonneau S, Whitehead V, Collin I, et al. The Montreal Cognitive Assessment, MoCA: a brief screening tool for mild cognitive impairment. J Am Geriatr Soc 2005;53(4):695-699. [doi: 10.1111/j.1532-5415.2005.53221.x] [Medline: 15817019]
12. Dumont L, Chénier-Leduc G, de Guise É, de Guinea AO, Sénécal S, Léger PM. Using a cognitive analysis grid to inform information systems design. In: Davis FD, Riedl R, vom Brocke J, Léger PM, Randolph AB, editors. Information Systems and Neuroscience: Gmunden Retreat on NeuroIS 2015. Cham: Springer International Publishing; 2015:193-199.
13. Weintraub S, Dikmen SS, Heaton RK, Tulsky DS, Zelazo PD, Bauer PJ, et al. Cognition assessment using the NIH Toolbox. Neurology 2013 Mar 12;80(11 Suppl 3):S54-S64 [FREE Full text] [doi: 10.1212/WNL.0b013e3182872ded] [Medline: 23479546]

XSL•FO
RenderX

14. Allen B. Cognitive abilities and information system usability. Inf Process Manage 1994;30(2):177-191. [doi: 10.1016/0306-4573(94)90063-9]
15. Wothke W, Curran LT, Augustin JW, Guerrero C, Bock RD, Fairbank BA, et al. Factor Analytic Examination of the Armed Services Vocational Aptitude Battery (ASVAB) and the KIT of Factor-Referenced Tests. San Antonio, TX: Operational Technologies Corp; 1991. URL: http://www.iapsych.com/wmfhcaarchive/LinkedDocuments/WOTH01.pdf [accessed 2024-04-26]
16. Oliveira J, Gamito P, Alghazzawi DM, Fardoun HM, Rosa PJ, Sousa T, et al. Performance on naturalistic virtual reality tasks depends on global cognitive functioning as assessed via traditional neurocognitive tests. Appl Neuropsychol Adult 2018;25(6):555-561. [doi: 10.1080/23279095.2017.1349661] [Medline: 28805447]
17. Czaja SJ, Charness N, Fisk AD, Hertzog C, Nair SN, Rogers WA, et al. Factors predicting the use of technology: findings from the Center for Research and Education on Aging and Technology Enhancement (CREATE). Psychol Aging 2006;21(2):333-352 [FREE Full text] [doi: 10.1037/0882-7974.21.2.333] [Medline: 16768579]
18. Khalili-Mahani N, Assadi A, Li K, Mirgholami M, Rivard ME, Benali H, et al. Reflective and reflexive stress responses of older adults to three gaming experiences in relation to their cognitive abilities: mixed methods crossover study. JMIR Ment Health 2020;7(3):e12388 [FREE Full text] [doi: 10.2196/12388] [Medline: 32213474]
19. Schatz R, Egger-Lampl S, Masuch K. The impact of test duration on user fatigue and reliability of subjective quality ratings. J Audio Eng Soc 2012;60(1):63-73.
20. Lallemand C, Gronier G. UX design methods: 30 fundamental methods to design for optimal experiences. Paris, France: Eyrolles; 2018.
21. Isaacs E, Oates J, ILSI Europe a.i.s.b.l. Nutrition and cognition: assessing cognitive abilities in children and young people. Eur J Nutr 2008;47(Suppl 3):4-24. [doi: 10.1007/s00394-008-3002-y] [Medline: 18683026]
22. Robbins TW, Sahakian BJ. Computer methods of assessment of cognitive function. In: Copeland JRM, Abou-Saleh MT, Blazer DG, editors. Principles and Practice of Geriatric Psychiatry. Chichester, West Sussex, England: Wiley; 2002:147-151.
23. Slegers K, Van Boxtel MPJ, Jolles J. The efficiency of using everyday technological devices by older adults: the role of cognitive functions. Ageing Soc 2009;29(2):309-325. [doi: 10.1017/s0144686x08007629]
24. Chan MY, Haber S, Drew LM, Park DC. Training older adults to use tablet computers: does it enhance cognitive function? Gerontologist 2016;56(3):475-484 [FREE Full text] [doi: 10.1093/geront/gnu057] [Medline: 24928557]
25. Kieras DE, Meyer DE. An overview of the EPIC architecture for cognition and performance with application to human-computer interaction. Hum–Comput Int 1997;12(4):391-438. [doi: 10.1207/s15327051hci1204_4]
26. Czaja SJ, Sharit J, Nair S, Rubert M. Understanding sources of user variability in computer-based data entry performance. Behav Inf Technol 1998;17(5):282-293. [doi: 10.1080/014492998119355]
27. Miyake A, Friedman NP, Emerson MJ, Witzki AH, Howerter A, Wager TD. The unity and diversity of executive functions and their contributions to complex "Frontal Lobe" tasks: a latent variable analysis. Cogn Psychol 2000;41(1):49-100. [doi: 10.1006/cogp.1999.0734] [Medline: 10945922]
28. Kiesel A, Steinhauser M, Wendt M, Falkenstein M, Jost K, Philipp AM, et al. Control and interference in task switching—a review. Psychol Bull 2010;136(5):849-874. [doi: 10.1037/a0019842] [Medline: 20804238]
29. Gade M, Schuch S, Druey M, Koch I. Inhibitory control in task switching. In: Grange J, Houghton G, editors. Task Switching and Cognitive Control. New York, NY: Oxford University Press; 2014:137-159.
30. US Army. Army individual test battery. In: Manual of Directions and Scoring. Washington, DC: War Department, Adjunct General's Office; 1944.
31. Lezak MD. Executive functions and motor performance. In: Neuropsychological Assessment. New York: Oxford University Press; 1995:650-685.
32. Mitrushina M, Boone KB, Razani J, D'Elia L. Handbook of Normative Data for Neuropsychological Assessment, Second Edition. New York, NY: Oxford University Press; 2005.
33. Reitan RM. Validity of the trail making test as an indicator of organic brain damage. Percept Mot Skills 2016;8(3):271-276. [doi: 10.2466/pms.1958.8.3.271]
34. Reitan RM, Wolfson D. Conventional intelligence measurements and neuropsychological concepts of adaptive abilities. J Clin Psychol 1992;48(4):521-529. [doi: 10.1002/1097-4679(199207)48:4<521::aid-jclp2270480414>3.0.co;2-c] [Medline: 1517447]
35. Strauss E, Sherman EMS, Spreen O. A Compendium of Neuropsychological Tests: Administration, Norms, and Commentary. Oxford: Oxford University Press; 2006.
36. Sánchez-Cubillo I, Periáñez JA, Adrover-Roig D, Rodríguez-Sánchez JM, Ríos-Lago M, Tirapu J, et al. Construct validity of the trail making test: role of task-switching, working memory, inhibition/interference control, and visuomotor abilities. J Int Neuropsychol Soc 2009;15(3):438-450. [doi: 10.1017/S1355617709090626] [Medline: 19402930]
37. Groff MG, Hubble LM. A factor analytic investigation of the Trail Making Test. Clin Neuropsychol 1981;3(4):11-13. [doi: 10.1007/978-0-387-79948-3_217]
38. Schear JM, Sato SD. Effects of visual acuity and visual motor speed and dexterity on cognitive test performance. Arch Clin Neuropsychol 1989;4(1):25-32. [doi: 10.1016/0887-6177(89)90004-8]

XSL•FO
RenderX

39. Gaudino EA, Geisler MW, Squires NK. Construct validity in the trail making test: what makes Part B harder? J Clin Exp Neuropsychol 1995;17(4):529-535. [doi: 10.1080/01688639508405143] [Medline: 7593473]
40. Crowe SF. The differential contribution of mental tracking, cognitive flexibility, visual search, and motor speed to performance on parts A and B of the trail making test. J Clin Psychol 1998;54(5):585-591. [doi: 10.1002/(sici)1097-4679(199808)54:5<585::aid-jclp4>3.0.co;2-k]
41. Meegan DV, Tipper SP. Visual search and target-directed action. J Exp Psychol Hum Percept Perform 1999;25(5):1347-1362. [doi: 10.1037//0096-1523.25.5.1347]
42. Treisman AM, Gelade G. A feature-integration theory of attention. Cogn Psychol 1980;12(1):97-136. [doi: 10.1016/0010-0285(80)90005-5] [Medline: 7351125]
43. Pietrzak RH, Maruff P, Mayes LC, Roman SA, Sosa JA, Snyder PJ. An examination of the construct validity and factor structure of the Groton Maze Learning Test, a new measure of spatial working memory, learning efficiency, and error monitoring. Arch Clin Neuropsychol 2008;23(4):433-445 [FREE Full text] [doi: 10.1016/j.acn.2008.03.002] [Medline: 18448309]
44. Gossweiler R, Kamvar M, Baluja S. What's up CAPTCHA? A CAPTCHA based on image orientation. New York, NY: Association for Computing Machinery; 2009 Presented at: Proceedings of the 18th international conference on World wide web; April 20-24, 2009; Madrid, Spain p. 841-850. [doi: 10.1145/1526709.1526822]
45. Kunda M, McGreggor K, Goel AK. A computational model for solving problems from the Raven's progressive matrices intelligence test using iconic visual representations. Cogn Syst Res 2013;22-23:47-66. [doi: 10.1016/j.cogsys.2012.08.001]
46. Aladwani AM, Palvia PC. Developing and validating an instrument for measuring user-perceived web quality. Inf Manag 2002;39(6):467-476. [doi: 10.1016/s0378-7206(01)00113-6]
47. Vignerons d'Exception. URL: https://www.vigneronsdexception.com/en/ [accessed 2024-06-03]
48. Renaud-Bray. URL: https://www.renaud-bray.com/accueil.aspx [accessed 2024-06-03]
49. LesPAC. URL: https://www.lespac.com/en/ [accessed 2024-06-05]
50. Lufa Farms. URL: https://montreal.lufa.com/en/#/ [accessed 2024-06-03]
51. Craigslist. URL: https://montreal.craigslist.org/ [accessed 2024-06-03]
52. Bowie CR, Harvey PD. Administration and interpretation of the Trail Making Test. Nat Protoc 2006;1(5):2277-2281. [doi: 10.1038/nprot.2006.390] [Medline: 17406468]
53. Thomas E, Snyder PJ, Pietrzak RH, Maruff P. Behavior at the choice point: decision making in hidden pathway maze learning. Neuropsychol Rev 2014;24(4):514-536. [doi: 10.1007/s11065-014-9272-7] [Medline: 25348261]
54. Gould MC, Perrin FAC. A comparison of the factors involved in the maze learning of human adults and children. J Exp Psychol 1916;1(2):122-154. [doi: 10.1037/h0072916]
55. Cuvillier M, Léger P, Sénécal S. Studies on the reliability, validity and sensitivity of single-item measurement scales in user experience. HEC Montréal 2021:57-93. [doi: 10.2307/j.ctv10qqx59.12]
56. Passalacqua M, Léger PM, Nacke LE, Fredette M, Labonté-Lemoyne E, Lin X, et al. Playing in the backstore: interface gamification increases warehousing workforce engagement. Ind Manag Data Syst 2020;120(7):1309-1330. [doi: 10.1108/imds-08-2019-0458]
57. Bracken MR, Mazur-Mosiewicz A, Glazek K. Trail making test: comparison of paper-and-pencil and electronic versions. Appl Neuropsychol Adult 2019;26(6):522-532. [doi: 10.1080/23279095.2018.1460371] [Medline: 30265569]
58. Tombaugh TN. Trail making test a and B: normative data stratified by age and education. Arch Clin Neuropsychol 2004;19(2):203-214 [FREE Full text] [doi: 10.1016/S0887-6177(03)00039-8] [Medline: 15010086]
59. Giovagnoli AR, Del Pesce M, Mascheroni S, Simoncelli M, Laiacona M, Capitani E. Trail making test: normative values from 287 normal adult controls. Ital J Neurol Sci 1996;17(4):305-309. [doi: 10.1007/BF01997792] [Medline: 8915764]
60. Arango-Lasprilla JC, Rivera D, Aguayo A, Rodríguez W, Garza MT, Saracho CP, et al. Trail making test: normative data for the Latin American Spanish speaking adult population. NeuroRehabilitation 2015;37(4):639-661 [FREE Full text] [doi: 10.3233/NRE-151284] [Medline: 26639932]
61. Shiffrin RM, Schneider W. Controlled and automatic human information processing: II. Perceptual learning, automatic attending and a general theory. Psychol Rev 1977;84(2):127-190. [doi: 10.1037//0033-295x.84.2.127]
62. Karran AJ, Demazure T, Leger P, Labonte-LeMoyne E, Senecal S, Fredette M, et al. Toward a hybrid passive BCI for the modulation of sustained attention using EEG and fNIRS. Front Hum Neurosci 2019;13:393 [FREE Full text] [doi: 10.3389/fnhum.2019.00393] [Medline: 31780914]
63. Guertin-Lahoud S, Coursaris CK, Sénécal S, Léger PM. User experience evaluation in shared interactive virtual reality. Cyberpsychol Behav Soc Netw 2023;26(4):263-272 [FREE Full text] [doi: 10.1089/cyber.2022.0261] [Medline: 37071640]
64. Courtemanche F, Labonté-LeMoyne E, Léger PM, Fredette M, Senecal S, Cameron A, et al. Texting while walking: an expensive switch cost. Accid Anal Prev 2019;127:1-8. [doi: 10.1016/j.aap.2019.02.022] [Medline: 30826692]
65. Giroux F, Couture L, Lasbareille C, Boasen J, Stagg CJ, Fleming MK, et al. Usability evaluation of assistive technology for ICT accessibility: lessons learned with stroke patients and able-bodied participants experiencing a motor dysfunction simulation. In: Davis FD, Riedl R, vom Brocke J, Léger PM, Randolph AB, Müller-Putz GR, editors. Information Systems and Neuroscience: NeuroIS Retreat 2022. Cham: Springer International Publishing; 2022:349-359.

XSL•FO
RenderX

66. Zolyomi A, Snyder J. Social-emotional-sensory design map for affective computing informed by neurodivergent experiences. Proc ACM Hum-Comput Interact 2021;5(CSCW1):1-37 [FREE Full text] [doi: 10.1145/3449151]
67. Demmans Epp C, McEwen R, Campigotto R, Moffatt K. Information practices and user interfaces: student use of an iOS application in special education. Educ Inf Technol 2015;21(5):1433-1456. [doi: 10.1007/s10639-015-9392-6]
68. Vasseur A, Léger PM, Courtemanche F, Labonte-Lemoyne E, Georges V, Valiquette A, et al. Distributed remote psychophysiological data collection for UX evaluation: a pilot project. In: Human-Computer Interaction. Theory, Methods and Tools: Thematic Area, HCI 2021, Held as Part of the 23rd HCI International Conference, HCII 2021, Virtual Event, July 24–29, 2021, Proceedings, Part I. Cham, Switzerland: Springer Nature; 2021.
69. Amershi S, Weld D, Vorvoreanu M, Fourney A, Nushi B, Collisson P, et al. Guidelines for human-AI interaction. New York, NY, United States: Association for Computing Machinery; 2019 Presented at: Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems; May 4-9, 2019; Glasgow, Scotland, UK p. 1-13. [doi: 10.1145/3290605.3300233]
70. Hussain J, Ul Hassan A, Muhammad Bilal HS, Ali R, Afzal M, Hussain S, et al. Model-based adaptive user interface based on context and user experience evaluation. J Multimodal User Interfaces 2018;12(1):1-16. [doi: 10.1007/s12193-018-0258-2]

Abbreviations

CAPTCHA: Completely Automated Public Turing tests to tell Computers and Humans Apart
MMSE: Mini-Mental State Examination
MoCA: Montreal Cognitive Assessment
NumRec: numerical recognition
PicRec: pictogram recognition
RM MANCOVA: repeated-measures multivariate analysis of covariance
RPM: Raven's Progressive Matrices
TMT: Trail Making Test
UX: user experience

Edited by A Kushniruk; submitted 22.06.23; peer-reviewed by L Loeb, C Baxter; comments to author 20.01.24; revised version received 13.02.24; accepted 20.02.24; published 14.06.24.

Please cite as:
Depauw T, Boasen J, Léger PM, Sénécal S
Assessing the Relationship Between Digital Trail Making Test Performance and IT Task Performance: Empirical Study
JMIR Hum Factors 2024;11:e49992
URL: https://humanfactors.jmir.org/2024/1/e49992
doi:10.2196/49992
PMID:38875007

©Tanguy Depauw, Jared Boasen, Pierre-Majorique Léger, Sylvain Sénécal. Originally published in JMIR Human Factors (https://humanfactors.jmir.org), 14.06.2024. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on https://humanfactors.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Physical Therapists' Acceptance of a Wearable, Fabric-Based Sensor System (Motion Tape) for Use in Clinical Practice: Qualitative Focus Group Study

Audrey Lee[1], MS; Patricia Dionicio[2], MPH; Emilia Farcas[3], PhD; Job Godino[3], PhD; Kevin Patrick[3,4], MD, MS; Elijah Wyckoff[5], MS; Kenneth J Loh[5], PhD; Sara Gombatto[6], PT, PhD

[1]Department of Bioengineering, San Diego State University, San Diego, CA, United States

[2]Joint Doctoral Program in Public Health, San Diego State University and University of California San Diego, San Diego, CA, United States

[3]Qualcomm Institute, University of California San Diego, La Jolla, CA, United States

[4]School of Public Health, University of California San Diego, La Jolla, CA, United States

[5]Active, Responsive, Multifunctional, and Ordered-materials Research (ARMOR) Laboratory, Department of Structural Engineering, University of California San Diego, La Jolla, CA, United States

[6]School of Exercise & Nutritional Sciences, College of Health & Human Services, San Diego State University, San Diego, CA, United States

Corresponding Author:
Sara Gombatto, PT, PhD
School of Exercise & Nutritional Sciences
College of Health & Human Services
San Diego State University
5500 Campanile Drive
San Diego, CA, 92182-7251
United States
Phone: 1 619 594 3288
Email: sgombatto@sdsu.edu

Abstract

Background: Low back pain (LBP) is a costly global health condition that affects individuals of all ages and genders. Physical therapy (PT) is a commonly used and effective intervention for the management of LBP and incorporates movement assessment and therapeutic exercise. A newly developed wearable, fabric-based sensor system, Motion Tape, uses novel sensing and data modeling to measure lumbar spine movements unobtrusively and thus offers potential benefits when used in conjunction with PT. However, physical therapists' acceptance of Motion Tape remains unexplored.

Objective: The primary aim of this research study was to evaluate physical therapists' acceptance of Motion Tape to be used for the management of LBP. The secondary aim was to explore physical therapists' recommendations for future device development.

Methods: Licensed physical therapists from the American Physical Therapy Association Academy of Leadership Technology Special Interest Group participated in this study. Overall, 2 focus groups (FGs; N=8) were conducted, in which participants were presented with Motion Tape samples and examples of app data output on a poster. Informed by the Technology Acceptance Model, we conducted semistructured FGs and explored the wearability, usefulness, and ease of use of and suggestions for improvements in Motion Tape for PT management of LBP. FG data were transcribed and analyzed using rapid qualitative analysis.

Results: Regarding wearability, participants perceived that Motion Tape would be able to adhere for several days, with some variability owing to external factors. Feedback was positive for the low-profile and universal fit, but discomfort owing to wires and potential friction with clothing was of concern. Other concerns included difficulty with self-application and potential skin sensitivity. Regarding usefulness, participants expressed that Motion Tape would enhance the efficiency and specificity of assessments and treatment. Regarding ease of use, participants stated that the app would be easy, but data management and challenges with interpretation were of concern. Physical therapists provided several recommendations for future design improvements including having a wireless system or removable wires, customizable sizes for the tape, and output including range of motion data and summary graphs and adding app features that consider patient input and context.

Conclusions: Several themes related to Motion Tape's wearability, usefulness, and ease of use were identified. Overall, physical therapists expressed acceptance of Motion Tape's potential for assessing and monitoring low back posture and movement, both

XSL•FO
RenderX

within and outside clinical settings. Participants expressed that Motion Tape would be a valuable tool for the personalized treatment of LBP but highlighted several future improvements needed for Motion Tape to be used in practice.

(JMIR Hum Factors 2024;11:e55246) doi:10.2196/55246

KEYWORDS

low back pain; physical therapy; physical therapist; wearable sensor; technology acceptance model; motion tape; kinesiology tape

Introduction

Prevalence and Impact of Low Back Pain

Low back pain (LBP) is one of the world's leading causes of disability [1-3]. In 2019, there were approximately 568.4 million prevalent cases, 223.5 million incident cases, and 63.7 million cases of years lived with disability owing to LBP reported globally [4]. LBP affects all ages and genders, but its prevalence increases with age, peaking at the age of approximately 45 to 54 years [4]. Approximately 70% to 85% of adults are expected to experience at least 1 episode of LBP in their lifetime [5]. Once predisposed to LBP, individuals are twice as likely to experience recurrent episodes of LBP [6]. Annually, LBP in the United States results in 149 million missed work days [7]. The total costs of LBP worldwide amount to approximately US $100 billion a year, with two-thirds of this amount owing to lost wages and decreased work productivity [8].

Treatment of LBP With Physical Therapy

Physical therapy (PT) is a common, effective, evidence-based treatment for LBP [9,10]. Specifically, active interventions including exercises prescribed by a physical therapist are effective for prevention and treatment of LBP [11,12]. During an initial examination, a physical therapist can identify musculoskeletal and neuromuscular impairments associated with the LBP problem by conducting assessment of the patient's posture and movement. Then, the physical therapist and patient can work together to promote strength, stability, and mobility with in-clinic sessions and an assigned home exercise program with the goal of decreasing pain and disability [10,13]. Monitoring the patient's posture and movement can provide a basis for determining individualized factors associated with the LBP problem, which can then be addressed through targeted interventions.

Incorporation of Technology in PT

Whether at home or at work, specific movement patterns that are performed repeatedly have been identified as a significant risk factor for the development and persistence of LBP [2,14,15]. These movement patterns of the low back region can be characterized by evaluating the angle, velocity, and acceleration [16] and can assist in LBP diagnosis, treatment, and prevention. There are several approaches to monitoring spine posture and movement. Generally, when conducting a PT examination, clinicians visually monitor posture and movement or use tools that measure the range of movement such as goniometers or inclinometers [17], but an alternative approach is to use technology to help better quantify the objective measures of spine posture and movement and offer potential benefits such as remote monitoring [16,18,19] while the patient is away from the clinic.

Technologies for Monitoring LBP

To date, existing technologies used to measure spine posture and movement in research and practice include optical motion capture, inertial measurement units (IMUs), and other wearable sensors [20-22]. Despite the variety of systems available, they generally present ≥1 limitation. Optical motion capture systems offer great precision and accuracy in monitoring human movement. However, their applications are limited owing to space needs, cost, and level of expertise needed. IMUs are portable devices that measure metrics such as acceleration and orientation [23] and include a variety of wearable sensors such as accelerometers, gyroscopes, and magnetometers, making them ideal for collecting data in a free-living environment. However, when used for monitoring human movement, IMUs have several limitations including decreased accuracy and precision for measuring slow movements [24,25], difficulty with measuring the axial plane movement accurately, inability to account for the multisegmented nature of the spine [26], and the need for multiple IMUs to triangulate posture and movement of a segment that can be cumbersome to the wearer [27].

Motion Tape

Owing to the limitations of existing sensor systems for measuring spine posture and movement, there is a need to explore new sensor innovations to address this issue. Ideally, such an approach would be wearable, unobtrusive, and usable in a clinical environment during PT sessions and in a person's natural environment to support home-based care. Another desired requirement would be high accuracy while collecting posture and movement data for a prolonged period.

Motion Tape, developed by Loh and Lin [28], is a disposable, self-adhesive skin-strain sensor system made using graphene nanosheets coated onto commercially available kinesiology tape (also known as K-Tape) [29-33]. Motion Tape has piezoresistive properties based on the deformation of the integrated graphene nanosheets in the tape that makes it sensitive to strain [33]. In previous studies, Motion Tape has demonstrated stable performance under cyclic strains [33,34]. In addition, the Motion Tape sensor system has been tested on human participants [33,34], displaying accuracy in measuring skin strains and angles across biceps, knees, shoulders, wrists, and various other body regions when compared with IMUs and skin strains estimated using optical motion capture systems [35]. Overall, Motion Tape offers noninvasive, comfortable, and practical skin-strain measurements and can comprehensively capture complex movements and muscle engagement, especially when applied as a network of sensors [35].

XSL•FO
RenderX

Motion Tape for a Low Back Use Case

When used for a low back use case, Motion Tape provides a means to capture the lumbar spine's multisegmental nature and multiplanar movements [36]. Motion Tape's low-profile and stretchable nature allows it to be worn throughout the day for all human shapes and sizes, and it could be suitable for use in an individual's natural environment with minimal interference to their daily activities. Motion Tape provides unique sensing streams that can be used in machine learning and artificial intelligence models to optimize inferences related to the management of LBP. Specifically, Motion Tape for a low back use case can address several key issues in a physical therapist's management of LBP, including the following: expanding on the level of detail available during the clinical assessment of posture and movement, assessing spinal posture and movement in a free-living environment, use for the promotion of engagement and adherence with and precise performance of a prescribed home exercise program, and using the patient's response to treatment to make informed decisions for future treatment or other patients [37]. Although there are several potential benefits that Motion Tape may add to personalized health care for LBP, the acceptability of Motion Tape among physical therapists has yet to be assessed.

Physical Therapists' Acceptance of Motion Tape

The success of this device is dependent on user acceptance or one's belief that the device will help them perform their work better (ie, perceived usefulness) and that the device's performance benefits outweigh the effort of using the device (ie, perceived ease of use) [38]. Thus, it is vital to understand physical therapists' perspectives about Motion Tape and their willingness to use it in their practice, to inform future developments and improvement of the technology.

Problem Statement

The primary aim of this research study was to evaluate physical therapists' acceptance of Motion Tape for the management of LBP. The secondary purpose was to explore physical therapists' current needs and recommendations regarding future development of Motion Tape.

Methods

Device Description and Stage of Development

In this study, licensed physical therapists evaluated a prototype of Motion Tape and examples of data streams from the app for a low back use case. The Motion Tape samples evaluated in this study included the Motion Tape sensor system with conductive wire leads connected to both sides of the sample (Figure 1).

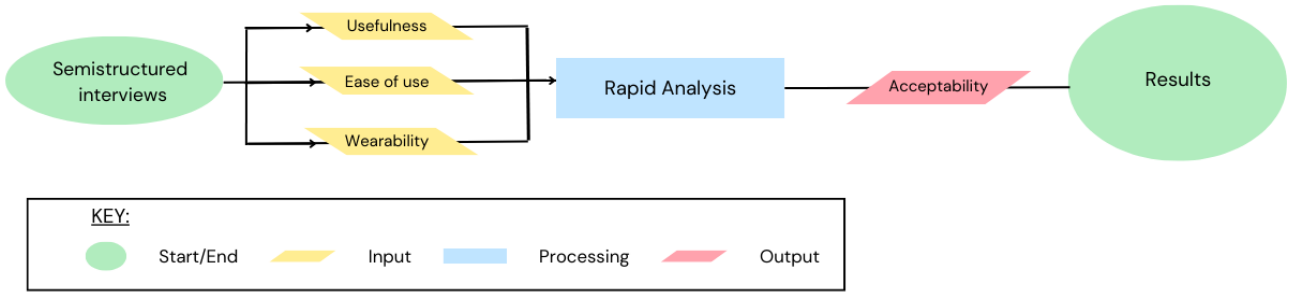
Figure 1. Motion Tape sample with conductive wire leads given to the physical therapists for evaluation.

Study Design

This exploratory, qualitative study was designed to explore physical therapists' acceptance of Motion Tape to provide a basis for future device development (Figure 2). The study was conducted from a constructivist point of view, with the goal of gaining insightful accounts and narrations of clinicians' lived experiences with technology and patients, rather than identifying an absolute truth [39]. We used semistructured focus groups (FGs) that incorporated human factor considerations to uncover real-world needs and obstacles and to ensure that the development of the sensor system can be informed by real-world PT clinical needs.

XSL•FO
RenderX

Figure 2. Study design overview—the evaluation of Motion Tape's acceptability.



Theoretical Framework and Constructs

The Technology Acceptance Model (TAM) framework was used in this study to assess two determinants of user acceptance of or willingness to use a technology: (1) perceived usefulness and (2) perceived ease of use [38,40]. An additional factor of wearability was also assessed to examine physical therapists' perceptions about patient-centered issues that would affect whether the device would be worn [41]. Recommendations for future improvements were also investigated to collect insight into data, device, and app developments that clinicians would like to see for Motion Tape.

> *Perceived usefulness was defined as the degree to which the use of Motion Tape would enhance the physical therapists' management of LBP [39-42], and this was assessed using the following constructs: (1) productivity, (2) effectiveness, (3) ability to make their job easy, and (4) benefits to PT treatment and recovery. Perceived ease of use was defined as the degree to which the use of Motion Tape would be effortless when used for managing LBP [39-42], and this was assessed using the following constructs: (1) how easy it would be for physical therapists to learn how to use it, (2) what level of instruction would physical therapists need to use it, and (3) how clear and understandable Motion Tape was in its current state. Wearability was defined as the degree to which Motion Tape would fit well and be comfortable for patients to wear on their back [42], and this was assessed based on (1) adhesion, (2) fit, (3) feel, and (4) how comfortable physical therapists would feel about applying and prescribing Motion Tape.*

Participants and Setting

This study was conducted at the American Physical Therapy Association's (APTA's) Combined Sections Meeting (San Diego, California) on February 24, 2023. Participants were recruited by sending study information via email to physical therapists who were members of the APTA Academy of Leadership Technology Special Interest Group. Members were also offered an opportunity to participate when they attended the Technology Special Interest Group in-person meeting at the APTA Combined Sections Meeting. Individuals were included in this study if they were a licensed physical therapist and were excluded from participating if they were unable to respond to questions in English. In total, 8 physical therapists were eligible and agreed to participate in 2 FGs of 4 clinicians each. A sample size of 8 people, in 2 FGs, was considered sufficient for this qualitative study to provide adequate variability and data saturation [43] and to provide a basis for device improvement. In addition, after data from the 2 FGs were collected and analyzed, the data were deemed saturated (ie, no new themes or codes were generated) and no further FGs were needed.

Ethical Considerations

The study protocol was considered to be exempt from ethics approval by the San Diego State University institutional review board. Each participant provided written consent before participating.

FG Methods

An FG guide (Multimedia Appendix 1) was used to lead the group's discussion. The FG guide was developed by investigators (AL, PD, and SG) to be semistructured with open-ended questions to explore the participants' perspectives about the usefulness, usability, and wearability of Motion Tape and to collect insight into future improvements for the sensors and data visualization (Textbox 1). A template of the FG guide was piloted with a Doctor of Physical Therapy student and a physical therapist at San Diego State University to ensure credibility [44]. General domains for each construct were prespecified to correspond with each interview question. Domains were defined based on the TAM framework and included perceived usefulness and perceived ease of use. An additional domain of wearability also was assessed.

XSL•FO
RenderX

Textbox 1. Guiding questions from the focus group guide.

Perceived wearability (W)

- How secure do you think the Motion Tape adhesive will be? (W-adhesion)
- To what degree do you think these sensors would fit your patients' anatomy (ie, their low back)? (W-fit)
- To what degree do you think your patients would feel the sensors on their back? (W-feel)
- How do you predict the Motion Tape Sensors would feel when being removed? (W-feel)

Perceived usefulness (U)

- To what degree would the usage of Motion Tape sensors affect how quickly you can assess your patient's posture, movement, or exercise performance? (U-efficiency)
- How effective do you think the Motion Tape sensors will be to capture valid data on your patients in the clinic? (U-effectiveness)
- How effective do you think the Motion Tape will be to capture valid data on your patients in their daily routine and normal environment? (U-effectiveness)
- To what degree would the usage of Motion Tape sensors affect the level of difficulty of your job as a clinician/physical therapist? (U–make job easier)
- What features, if any, would make the Motion Tape more useful to you? (U-useful)

Perceived ease of use (EU)

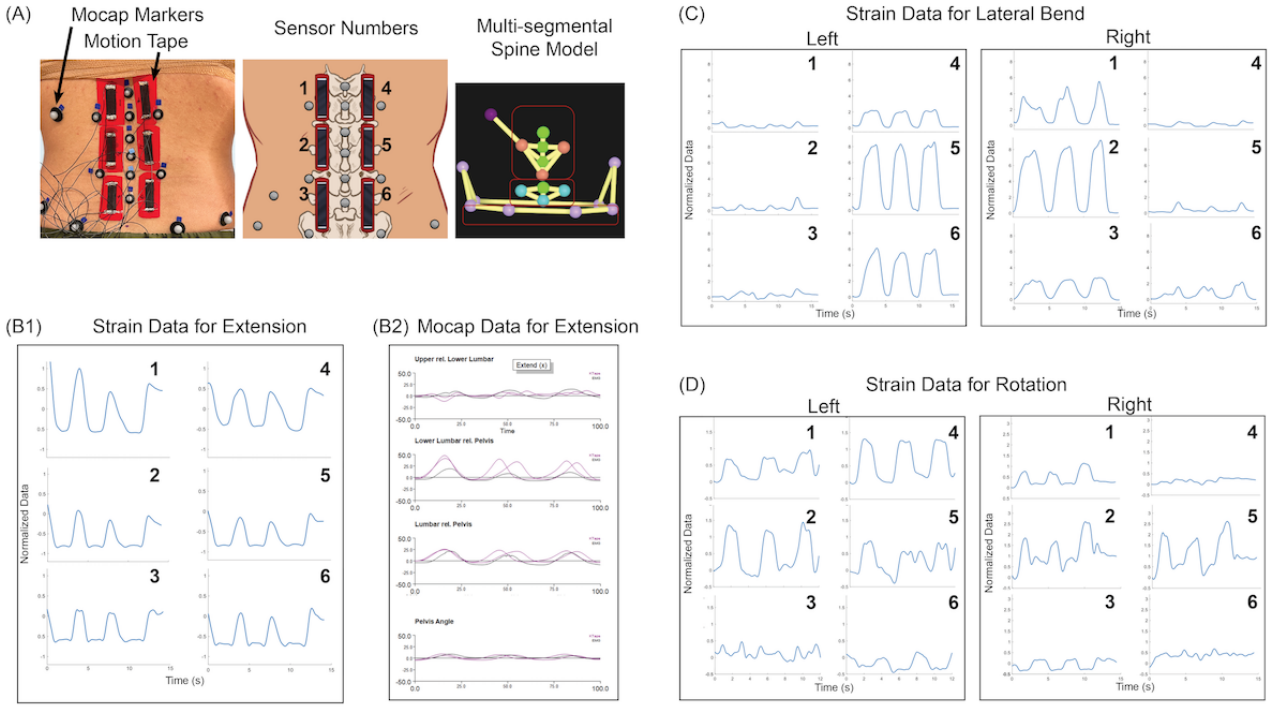
- How easy do you think it would be to learn how to use Motion Tape? (EU-easy to learn)
- How comfortable would you feel prescribing Motion Tape to a patient to monitor their movements at home? (EU-comfort in usage)
- What level of knowledge do you think a clinician/PT would need to use the Motion Tape? (EU-clear and understandable)
- How easy/difficult do you think it would be for a clinician/PT to apply the Motion Tape to the patient's back? (EU-easy to use)
- What features, if any, would make the Motion Tape easier for you to use? (EU-easy to use)

FGs were conducted by AL (a female Master of Science student investigator) and PD (a female PhD student investigator). Reflexivity was maintained by the research team by discussing assumptions and biases that may influence how the clinicians responded to the FG moderators, who were not licensed physical therapists. As SG is a licensed physical therapist and member of APTA, she was able to provide valuable insight during the development of the interview guide, analysis, and interpretation to ensure credibility of the findings [44].

FGs were anonymized, and each participant was assigned a color as a name to ensure confidentiality. Each FG lasted approximately 1 hour and was recorded using digital voice recorders (Olympus Voice Recorder; WS-853). Before asking the participants questions, the investigators gave each participant a sample of Motion Tape. Participants were then oriented to a poster that displayed the Motion Tape placement and app data output streams (Figure 3).

XSL•FO
RenderX

Figure 3. Poster of Motion Tape placement and app data output for a low back use case. (A) The laboratory setup with 6 pieces of Motion Tape and several optical motion capture markers on anatomical landmarks of the lumbar spine. (B) The graphs display the following: (1) blue—the normalized strain data for extension, captured by the 6 Motion Tapes, and (2) purple—the kinematics for extension in degrees, captured by the optical motion capture system (reference standard). (C) The normalized strain data for right and left lateral bending obtained from the 6 Motion Tapes. (D) The normalized strain data for right and left rotation obtained from the 6 Motion Tapes.



Data Processing and Analysis

All FG audio data were downloaded to a HIPAA (Health Insurance Portability and Accountability Act)-compliant laboratory server, accessible only to the research staff, and removed from the digital voice recorder. The recordings were then transcribed, first using computer-based transcription (Word; Microsoft Corp). An investigator then checked and verified each transcription by listening to the original audio and reviewing and correcting the computer-based transcription.

Considering the need for timely feedback in the sensor development process, we adopted a rapid qualitative analysis (RQA) approach to explore themes regarding the acceptability and wearability of Motion Tape [45]. RQA was conducted by 3 investigators to assess the FG responses effectively and efficiently and to identify major themes. Codes and themes for RQA were deductively developed based on the TAM framework and the study objective [41]. We then used an inductive approach to generate RQA codes and themes, allowing for quick sorting of FG dialogue.

To ensure rigor and consistency of the method, a constant comparative approach with investigator triangulation was used at each stage [46]. First, the 3 investigators independently completed a summary report for each FG, with quotes and relevant topics under the respective themes and codes. Once the individual coding and summary reports for both FGs were completed, the investigators consolidated them into a combined rapid analysis summary report for each FG, unifying themes and reconciling discrepancies by consensus through discussion. The summary reports for each FG were then transferred into a matrix in which each row was a participant quote and each column was a domain. From this matrix, investigators identified the underlying themes and subthemes between the 2 FGs.

Results

Overview

In total, 8 physical therapists (n=5, 63% men and n=3, 38% women), with a mean age of 47.5 (SD 5.6) years participated in this study. Participants reported obtaining PT degrees ranging from a bachelor's degree to a Doctorate in Physical Therapy and had, on average, 20 (SD 8.5) years of clinical practice experience, and most reported practicing in an outpatient orthopedic setting. Of the 8 participants, 5 (63%) reported having advanced doctoral degrees (3/5, 60% PhD; 2/5, 40% EdD).

The qualitative results from the FGs were organized using the TAM for the acceptance of Motion Tape [38,40-42]. Data were organized based on the 3 main domains relevant to user acceptance (perceived wearability, perceived usefulness, and perceived ease of use) and 21 subthemes (Textbox 2). Subthemes were further designated using *positive*, *negative*, and *neutral* valences. Positive valence indicates that the FG participants perceived the Motion Tape attribute as positive. Negative valence indicates that the FG participants perceived the attribute as negative. Neutral valence indicates that the FG participants perceived the attribute as neither positive nor negative.

XSL•FO
RenderX

Textbox 2. Themes (n=3), subthemes (n=21), and valences of user acceptance of Motion Tape.

Theme 1: perceived wearability

- Positive
  - Motion Tape has a small, universal fit.
  - The feeling of Motion Tape on the skin would decrease over time.
- Negative
  - Patients may feel Motion Tape's wires snagging or sensors rubbing on clothes.
  - Motion Tape does not consider people with skin sensitivities.
- Neutral
  - Motion Tape adheres for 3-4 days but may adhere less owing to external factors.
  - The feeling of Motion Tape being removed depends on the physical therapist.

Theme 2: perceived usefulness

- Positive
  - Motion Tape could increase specificity of physical therapy management of low back pain (LBP).
  - Motion Tape could be effective for the diagnosis, management, and monitoring of low back pain (LBP).
  - The feeling of Motion Tape and the awareness of Motion Tape monitoring would increase adherence to a home exercise program.
  - Motion Tape would be beneficial in telerehabilitation and hybrid sessions.
  - Motion Tape could increase the physical therapist's awareness of the pain source.
- Negative
  - Motion Tape brings legal concerns with data responsibility.
  - Motion Tape's reliability could be affected by external factors.
- Neutral
  - Motion Tape could increase the efficiency of assessments, but set up could take more time.

Theme 3: perceived ease of use

- Positive
  - Motion Tape would be easy for a physical therapist to apply.
- Negative
  - Motion Tape has a lot of data to sift through.
  - Motion Tape data are hard to interpret in their current state.
  - The self-application of Motion Tape would be difficult.
  - Motion Tape is designed for single use.
- Neutral
  - The prescription of Motion Tape is subjective to many factors.
  - The user interface would dictate how much knowledge would be needed to use Motion Tape.

Domain 1: Perceived Wearability

Regarding perceived *wearability*, all physical therapists were familiar with commercially available kinesiology tape. Thus, their thoughts about perceived wearability reflected their experience with kinesiology tape. For example, the physical therapists expected Motion Tape to last about 3 to 4 days. A physical therapist mentioned the following:

> *Oh, I've used the K-Tape for four days before it started peeling off. Sometimes it lasts more than five days actually. Three to four days I think is average.* [FG1]

XSL•FO
RenderX

However, some physical therapists clarified that the *longevity* of Motion Tape's adhesion depends on several factors. For example, 2 of the physical therapists expressed the following:

> *How secure it is depends on a lot of factors, like moisture on the skin. It depends on not just moisture, but how clean your skin is and how much hair is on the skin.* [FG2]
>
> *Some of them, specifically on the low back, tend to have more oily skin, and that depreciates the life of the tape.* [FG1]

Regarding the *fit* aspect of wearability, physical therapists also believed that Motion Tape's size was sufficiently small to be universal to the wearer and the placement location. They expressed the following:

> *In my experience with tapes like this, it fits most of the clientele that I've worked with, both inpatient and long-term post-acute.* [FG1]
>
> *If it was that little strip, I think it would be great to use anywhere.* [FG2]

Regarding the *feel* aspect of wearability, generally, physical therapists felt that patients would feel Motion Tape at first when applied but would become less aware over time until the tape starts to peel off:

> *They'd know that they're there, and they'd probably become less aware of over time.* [FG1]

However, physical therapists generally felt that with Motion Tape's current design, patients would *feel* the wires snagging or the sensors rubbing on clothing. A physical therapist explained it as follows:

> *So contraptions with wires will always have that uncomfortable feeling. Always. But if you go the wireless route, then probably after two days, the patient will be more comfortable until the tape starts peeling off. However, what I'm wearing right now, something that goes above my PSIS, if I go to the bathroom or do something, I'm going to, it's gonna move around, it might get pulled on it by my clothes.* [FG1]

When removing Motion Tape, physical therapists said that patient feelings about the removal process would be quite variable. Some physical therapists felt that it was subjective to how the therapist removed Motion Tape and how much hair or oil the individual has on their skin. A physical therapist explained it as follows:

> *I'm just thinking of whoever is taking it off. You know, like, it depends on you, like, some people just rip. And some people are just gentle. So subjective. So it depends on the training of the therapist and concern if they're empathetic to our patients.* [FG1]

The physical therapists mentioned some wearability concerns during the FGs. A concern was about how patients with skin sensitivities would be able to use Motion Tape. A physical therapist asked the following:

> *For those with skin allergy. Can you put an under wrap under this?* [FG1]

Domain 2: Perceived Usefulness

Physical therapists expressed mixed feelings about whether Motion Tape would increase their efficiency with assessments of lumbar spine posture and movement. Some expressed that if all they had to do was apply the tape, then there would be increased efficiency:

> *If it's easy to objectively document, by understanding the graph, I think it's a night and day difference versus getting into the goniometers and doing manual assessment. Instead, you put on the tape, ask the patient to rotate their trunk, lean forward, reach forward, extend their back. And then if I have it digitally by email or direct messaging, it would save a bunch of time.* [FG1]

However, others felt that it would reduce efficiency. A physical therapist explained the following:

> *Regarding the speed of assessment, I would be a little doubtful. I think by the time that you took this and you put it on the patient, you hooked up all the wires to it, you did the calibration, if you need to do a calibration, it might take just as long as doing an assessment. I would have concerns around the accuracy of this, to give you a number, an accurate range of motion, particularly for things like rotation. But if the data was convincing that everyone, if it was validated for everyone that gave you an accurate number, I think it could improve the quality of assessment.* [FG1]

A physical therapist felt that for the in-clinic assessments, Motion Tape would improve specificity:

> *I don't feel like it [Motion Tape] would improve speed, it would improve specificity.* [FG2]

Physical therapists also mentioned that they could envision Motion Tape as a useful tool for self-management and remote monitoring when used in combination with in-clinic PT. A physical therapist mentioned that the ability to monitor patients outside the clinic would be very meaningful:

> *That's the best place to actually observe them, their normal environment. If they're in therapy, they're being observed, coached, cued by a skilled clinician. Their performance is definitely going to be different. So if they're at home, and we're able to monitor them at home, I think the treatment will be more, and your adjustment and progression will be more meaningful.* [FG1]

Some physical therapists suggested that having patients wear these sensors would increase their awareness of being monitored and thus increase engagement with and adherence to the home exercise program:

> *I think that what it has to offer is improving...adherence with our programs. I think that's your potential.* [FG1]

XSL•FO
RenderX

> *When you tell someone, I'm looking at your posture right now, you change [gesturing to posture]. If they think you are watching, they'll do better.* [FG1]

Physical therapists expected patients to have a phenomenal experience with Motion Tape when used in a hybrid setting:

> *I think to his point that if it's applied properly in the clinics, it's hybridized, and you can take a call, and there's no technical involvement on the patient side, and all they do is open up the app, they'd have this really phenomenal experience.* [FG2]

Specifically, several physical therapists expressed that Motion Tape would help with the identification of postures and movements in free-living environments that provoke pain, allowing for more meaningful interventions:

> *I think for it to be very useful. It would have to compare with the app where you've got user input as to what's going on...where he's got these flags and the data that was pain here, pain here, pain here, and you can look, you know, to the periods of time before that.* [FG2]

Some physical therapists did have some concerns about the usefulness of Motion Tape. A physical therapist expressed legal concerns regarding data responsibility:

> *As long as you collect data, someone's then responsible for it. So who's going to look at it? What's the liability then that person takes on by having that information?...if something goes wrong, and the therapist hasn't looked at the data, I'd like to know, are they liable?* [FG1]

Another concern was knowing what external factors affect Motion Tape's signal and data reliability, mentioning that the use of Motion Tape in practice was "gonna depend on the reliability of the data" (FG2).

Several physical therapists felt that there were a variety of variables that might affect the reliability of the signal or data. They expressed the following:

> *And what other factors affect them, the sensors, as far as humidity, water, other environmental factors that might affect it? You know, what if they have a compression garment around the trunk, for example, does that affect the sensors?* [FG2]
>
> *Whether, getting it wet and getting so some things on it changes the conductivity, and therefore the calibration over time.* [FG2]
>
> *You get variability in the readings based on amount of tension that people put on it when they applied it.* [FG2]
>
> *Whether that's different from person to person because of different makeups morphology.* [FG1]

Domain 3: Perceived Ease of Use

Physical therapists felt that it would be easy to apply Motion Tape, given their background knowledge in human anatomy. A physical therapist stated the following:

> *You would need to know basic clinical knowledge of the application for where to look for the muscles, you know, right. So, they need to be clinician to have knowledge of the body.* [FG2]

When asked whether they would feel comfortable using Motion Tape with their patients, there were mixed responses among physical therapists. Some mentioned that it would depend on "cost and buy-in" (FG2) or how it was going to be "incorporated into the plan of care" (FG1). A physical therapist even explained the variability as follows:

> *Depends on the situation, honestly. I mean, I have some families that I'll show them how to do the application. And I'll see them three weeks later, and they've reapplied four times and done it great. And then I've seen others that I'm like, "Oh, no! This is nope."* [FG2]

There were also several concerns about the ease of use. Some physical therapists felt that they would have challenges with ease of use, specifically regarding interpretation of the data:

> *I think in its current form, easy to apply. Hard to interpret.* [FG1]
>
> *It depends on the interface and how much it interprets the points. The tape will be easy, but it's all the other pieces.* [FG2]

Additional concerns about the ease of use included that the amount of data presented was excessive and the type of data displayed was difficult to interpret. The physical therapists expressed a desire to see the range of motion displayed in degrees rather than resistance in ohms:

> *I think I'm probably realistically just correlated with what they report has been painful. Because I don't know that I've ever been so interested in all of that. Like, it might be too much data. For a patient, like I don't necessarily need to know their range of motion during every single activity, I need to know when it is relevant to them. And when it is impacting whatever condition they're here for.* [FG1]
>
> *And again, I think for a clinician, it's going to have to be meaningful data. It's gonna have be Range of motion data not ohms.* [FG1]
>
> *So then, conceivably, would it be helpful instead of giving you normalized strain,...if they could interpret it, would convert this over to degrees of rotation and flexibility?* [FG2]
>
> *If you could get range of motion kind of information, I think that would be great.* [FG2]

Another concern was about how challenging the self-application would be for patients:

> *How are people actually going to apply this on their own, someone that doesn't know how?* [FG2]

Finally, another concern was that Motion Tape is a single-use product. A physical therapist explained the challenge of a single-use product as follows:

XSL•FO
RenderX

Okay, now how about waste? So it's like a single use thing? Now I'm gonna throw in a whole planet into this is single usage. Or can you reapply? [FG2]

Future Recommendations

Future recommendations from the physical therapists were organized into 3 categories (Textbox 3): data, physical features, and app features.

Textbox 3. Themes (n=3) and subthemes of future recommendations for Motion Tape.

Theme 1: Data recommendations

- Motion Tape data should be easy to read at a glance.
- Motion tape data should account for differing patient morphology.
- Physical therapists should be aware of factors that affect Motion Tape data.

Theme 2: Physical feature recommendations

- Motion Tape should be made wireless or with removable wires.
- Motion Tape should be reusable.
- Motion Tape should be customizable in length.

Theme 3: App feature recommendations

- Motion Tape app should include BMI input.
- Motion Tape app should include input for a patient's change in activity.
- Motion Tape app should allow flagging events.
- Motion Tape app should include comparative data.

Regarding the data *recommendations*, physical therapists expressed that data should be summarized in the form of an at-a-glance graph with 1 overall meaningful number, reflecting the range of motion. They would also like to know how the data change from person to person owing to morphology and how external factors (water, application stretch variability, and skin movement) affect the data. Additional data that would be useful for their job included comparative data, graphs with a color scale, and information about muscle activation. Participants in an FG expressed the following:

Take a baseline and have them rotate from that position and determined by the volume strain, whether they are tension either degrees, or even if it's yellow, green, yellow, red, like if they're moving within if they can't pinpoint it specifically, but you know, within a range, would that be helpful? [FG2]

I think even just having comparative data would be helpful, right? Because, you know, I keep telling my students, "Don't tell me, 'I want to increase range 10 degrees.'" Because that doesn't tell me, "Can they walk?" Right? But, "Are they doing better now than they were doing when we started?" That's useful. So even if we get baseline data that could be translated into amount of motion and then follow up data that says, "hey, it's more, it's more fluid, it's better, it's whatever." I think that can be really useful. Now I know the payers are gonna want, how much rotation did you get? How much lateral flexion did you get? [FG2]

And I think beyond the range of motion, I work in neuro. I think just like muscle activation would be interesting, you know, like, how much activation did you get today, for example, versus six weeks ago, post stroke or, you know, spinal cord or something? I think that would be really interesting just to see the firing muscle activation. And on the flip side, and I don't know if that's possible, but looking at specificity. Could that be something to monitor changes in specimen specificity? Post- X Y & Z intervention, right? That could also be interesting. So it's not really about range of motion, we're also activity known as firing or not? [FG2]

Regarding *physical feature* recommendations, physical therapists wanted a way to mitigate the wires, either by moving to a wireless system or making them removable. Physical therapists were also concerned about the limited stretchability of the short pieces of tape, as it would not be long enough in length for typical kinesiology tape use, and recommended making the length customizable to the physical therapists' needs. Physical therapists were also concerned about Motion Tape's single-use design and were curious about whether it could be reusable to reduce waste:

Again, I'm thinking like, in the future, no wires, you've got a strip of graphene that you could customize length to, with those couple millimeters around the edge. And if we wanted a whole length, we cut whole lengths. And if we want segments, we can cut segments. And it feeds the data to the app somehow tailor it to someone's body. [FG2]

So you can imagine that maybe something like this could be a roll of tape. Yeah, the width of duct tape. And there's actually two pieces on this roll. There's one section, that's the conductive piece, that you can cut it to length, and then next to it there are maybe

XSL•FO
RenderX

> *there's a wire section, that's conductive tape that you can pull off and put on the ends of whatever you choose. So you get one roll of tape. And then one of them is the is the graphene is the other piece that you tear off to the appropriate length is the conductive tape that connects it to the box. And then it's a solution, you can customize length and you have your conductive piece and then your graphene.* [FG2]

Regarding future *app feature* recommendations, physical therapists expressed a need for the capability to input factors such as BMI, activity changes, "flags" for events, and changes in pain to help label, compare, or contextualize the data.

Discussion

Overview

There is a gap in the research between rehabilitation device development and evaluation of clinicians' acceptance of such devices. Most existing studies have considered patient or user satisfaction [47,48], whereas others that consider the clinician's perspective have not specifically evaluated sensors for measuring spine posture and movement [49,50]. In this study, several themes relating to physical therapists' perspectives about Motion Tape's wearability, usefulness, and ease of use for a low back use case were identified.

Domain 1: Perceived Wearability

One of the most common challenges for wearable sensors is ensuring that they are unobtrusive to the wearer's natural movement and environment [39]. The small form and fit of Motion Tape was considered by physical therapists to be ideal for a wearable sensor. However, similar to previous studies, the wires in the current design were considered to be not ideal [37]. Studies have shown that wireless technologies tend to be more widely used in many fields, especially in the field of wearable devices for health care [51]. Thus, a future iteration of Motion Tape without wires would be considered optimal. On the basis of feedback obtained from physical therapists, wearability for people with skin sensitivities also should be considered. Previous studies have shown that skin irritation is the most common concern when using kinesiology tape for extended periods of time [52,53]. Thus, future studies should explore whether a medium or substrate can be used under Motion Tape to mitigate skin irritation, possibly as an extension of recent research that integrated Motion Tape with elastic fabric for respiration monitoring [54].

Domain 2: Perceived Usefulness

There were mixed feelings among physical therapist participants about how efficient Motion Tape would be in the clinic. Overall, most physical therapists felt that Motion Tape would increase the specificity of their assessments, a characteristic that has been shown to be beneficial for LBP diagnosis and treatment [55]. Furthermore, Motion Tape's ability to monitor the patient's movements remotely was considered beneficial, as this feature may increase adherence to home exercise programs, which is an important component of effective treatment for LBP [56,57].

Domain 3: Perceived Ease of Use

On the basis of physical therapists' perspectives, Motion Tape would be easy to apply, but data would be difficult to interpret. Creating a device that is easy to use and understand is crucial because it predicts consumer use behavior [38,41]. Recommendations included presenting the data in units that physical therapists are more familiar with (ie, degrees of range of motion) and creating an app that requires minimal time for the physical therapists to use. These changes may promote increased device use and acceptance in PT.

Future Recommendations

On the basis of clinician feedback, Motion Tape appears to be a promising new technology that could be used for monitoring lumbar spine posture and movement in the management of patients with LBP. Future device development will be needed to address clinician recommendations obtained from this study in the domains of wearability and ease of use. In addition, future studies will be needed to validate Motion Tape in laboratory, clinical, and free-living environments and to investigate patient acceptance of Motion Tape.

Limitations

A limitation of this study is that participants were physical therapists who were part of a Technology Special Interest Group and are likely to be more receptive to using technology in practice. Thus, this study's results regarding Motion Tape's acceptability may be biased in favor of Motion Tape's ease of use, usefulness, and wearability. Future studies should also assess the acceptability of Motion Tape for clinicians who do not regularly use technology in their practice. Another limitation is that the physical therapists were not presented with active samples of Motion Tape with live data streams in the app. Instead, participants were given inactive samples of Motion Tape and presented with a poster with examples of app data streams. Future studies should provide an opportunity for physical therapists to apply Motion Tape to a person and use it with the app interface. Finally, there was a potential for investigator bias in the interpretation of the results, as several investigators of this study are actively working on the development of this device. However, 2 of the 3 investigators who conducted data analysis were outside the primary research team.

Conclusions

Physical therapists expressed overall acceptance of Motion Tape for its potential to monitor and assess low back posture and movement, both within and outside clinical settings. Physical therapist participants expressed that Motion Tape would be a valuable tool for personalized treatment of LBP but highlighted several future improvements needed for Motion Tape to be used in practice.

XSL•FO
RenderX

Acknowledgments

The authors would like to thank Nicolette Jaghab for her assistance with qualitative data analysis. This study is part of the Multi-Sensor Adaptive Data Analytics for Physical Therapy project and was supported by US National Science Foundation (award IIS-2205093).

Conflicts of Interest

KJL is a cofounder of JAK Labs Inc, a company that may potentially benefit from the study results. JAK Labs intends to commercialize Motion Tape for the physical therapy and rehabilitation market, among other markets. The terms of this arrangement have been reviewed and approved by the University of California San Diego in accordance with its conflicts of interest policies.

Multimedia Appendix 1
Focus group guide.
[DOCX File , 21 KB - humanfactors_v11i1e55246_app1.docx]

References

1. Hartvigsen J, Hancock MJ, Kongsted A, Louw Q, Ferreira ML, Genevay S, et al. What low back pain is and why we need to pay attention. Lancet 2018 Jul 09;391(10137):2356-2367 [FREE Full text] [doi: 10.1016/S0140-6736(18)30480-X] [Medline: 29573870]
2. Silva FG, da Silva TM, Maher C, L Ferreira M, Costa LD, Costa LO. What triggers an episode of acute low back pain? A protocol of a replication case-crossover study. BMJ Open 2021 Mar 05;11(2):e040784 [FREE Full text] [doi: 10.1136/bmjopen-2020-040784] [Medline: 33550235]
3. Wu A, March L, Zheng X, Huang J, Wang X, Zhao J, et al. Global low back pain prevalence and years lived with disability from 1990 to 2017: estimates from the Global Burden of Disease Study 2017. Ann Transl Med 2020 Mar;8(6):299 [FREE Full text] [doi: 10.21037/atm.2020.02.175] [Medline: 32355743]
4. Chen S, Chen M, Wu X, Lin S, Tao C, Cao H, et al. Global, regional and national burden of low back pain 1990-2019: a systematic analysis of the Global Burden of Disease study 2019. J Orthop Translat 2022 Jan;32:49-58 [FREE Full text] [doi: 10.1016/j.jot.2021.07.005] [Medline: 34934626]
5. Henchoz Y, Kai-Lik So A. Exercise and nonspecific low back pain: a literature review. Joint Bone Spine 2008 Oct;75(5):533-539. [doi: 10.1016/j.jbspin.2008.03.003] [Medline: 18801686]
6. Hestbaek L, Leboeuf-Yde C, Manniche C. Low back pain: what is the long-term course? A review of studies of general patient populations. Eur Spine J 2003 May;12(2):149-165 [FREE Full text] [doi: 10.1007/s00586-002-0508-5] [Medline: 12709853]
7. Kahere M, Hlongwa M, Ginindza TG. A scoping review on the epidemiology of chronic low back pain among adults in sub-Saharan Africa. Int J Environ Res Public Health 2022 Mar 03;19(5):2964 [FREE Full text] [doi: 10.3390/ijerph19052964] [Medline: 35270657]
8. Katz JN. Lumbar disc disorders and low-back pain: socioeconomic factors and consequences. J Bone Joint Surg Am 2006 May;88 Suppl 2:21-24. [doi: 10.2106/JBJS.E.01273] [Medline: 16595438]
9. Anar S. The effectiveness of home-based exercise programs for low back pain patients. J Phys Ther Sci 2016 Oct;28(10):2727-2730 [FREE Full text] [doi: 10.1589/jpts.28.2727] [Medline: 27821923]
10. Marich AV, Lanier VM, Salsich GB, Lang CE, Van Dillen LR. Immediate effects of a single session of motor skill training on the lumbar movement pattern during a functional activity in people with low back pain: a repeated-measures study. Phys Ther 2018 Jul 01;98(7):605-615 [FREE Full text] [doi: 10.1093/ptj/pzy044] [Medline: 29660077]
11. Saragiotto BT, Maher CG, Yamato TP, Costa LO, Menezes Costa LC, Ostelo RW, et al. Motor control exercise for chronic non-specific low-back pain. Cochrane Database Syst Rev 2016 Jan 08;2016(1):CD012004 [FREE Full text] [doi: 10.1002/14651858.CD012004] [Medline: 26742533]
12. Steffens D, Maher CG, Pereira LS, Stevens ML, Oliveira VC, Chapple M, et al. Prevention of low back pain: a systematic review and meta-analysis. JAMA Intern Med 2016 Mar;176(2):199-208. [doi: 10.1001/jamainternmed.2015.7431] [Medline: 26752509]
13. Nelson BW, O'Reilly E, Miller M, Hogan M, Wegner JA, Kelly C. The clinical effects of intensive, specific exercise on chronic low back pain: a controlled study of 895 consecutive patients with 1-year follow up. Orthopedics 1995 Oct;18(10):971-981. [doi: 10.3928/0147-7447-19951001-05] [Medline: 8584467]
14. Steffens D, Ferreira ML, Latimer J, Ferreira PH, Koes BW, Blyth F, et al. What triggers an episode of acute low back pain? A case-crossover study. Arthritis Care Res (Hoboken) 2015 Mar;67(3):403-410 [FREE Full text] [doi: 10.1002/acr.22533] [Medline: 25665074]
15. Karayannis NV, Jull GA, Hodges PW. Physiotherapy movement based classification approaches to low back pain: comparison of subgroups through review and developer/expert survey. BMC Musculoskelet Disord 2012 Mar 20;13:24 [FREE Full text] [doi: 10.1186/1471-2474-13-24] [Medline: 22348236]

XSL•FO
RenderX

16. Laird RA, Gilbert J, Kent P, Keating JL. Comparing lumbo-pelvic kinematics in people with and without back pain: a systematic review and meta-analysis. BMC Musculoskelet Disord 2014 Jul 10;15(1):229 [FREE Full text] [doi: 10.1186/1471-2474-15-229] [Medline: 25012528]
17. Zheng Z, Wang Y, Wang T, Wu Y, Li Y. A systematic review and meta-analysis on comparative kinematics in the lumbopelvic region in the patients suffering from spinal pain. J Healthc Eng 2022;2022:7369242 [FREE Full text] [doi: 10.1155/2022/7369242] [Medline: 35345658]
18. Alschuler KN, Hoodin F, Murphy SL, Geisser ME. Ambulatory monitoring as a measure of disability in chronic low back pain populations. Clin J Pain 2011 Oct;27(8):707-715. [doi: 10.1097/AJP.0b013e318217b7d0] [Medline: 21505322]
19. Schaller A, Rudolf K, Dejonghe L, Grieben C, Froboese I. Influencing factors on the overestimation of self-reported physical activity: a cross-sectional analysis of low back pain patients and healthy controls. Biomed Res Int 2016;2016:1497213 [FREE Full text] [doi: 10.1155/2016/1497213] [Medline: 27298820]
20. Shum GL, Crosbie J, Lee RY. Symptomatic and asymptomatic movement coordination of the lumbar spine and hip during an everyday activity. Spine (Phila Pa 1976) 2005 Dec 01;30(23):E697-E702. [doi: 10.1097/01.brs.0000188255.10759.7a] [Medline: 16319739]
21. Shum GL, Crosbie J, Lee RY. Effect of low back pain on the kinematics and joint coordination of the lumbar spine and hip during sit-to-stand and stand-to-sit. Spine (Phila Pa 1976) 2005 Oct 01;30(17):1998-2004. [doi: 10.1097/01.brs.0000176195.16128.27] [Medline: 16135992]
22. Bauer CM, Rast FM, Ernst MJ, Kool J, Luomajoki H, Suni J, et al. Validity and reliability of inertial measurement units when measuring lumbar range of motion, movement control, repetetive movement and reposition error. Physiotherapy 2015 May;101:e914-e915. [doi: 10.1016/j.physio.2015.03.1754]
23. Victorino MN, Jiang X, Menon C. Chapter 7 - wearable technologies and force myography for healthcare. In: Tong RK, editor. Wearable Technology in Medicine and Health Care. Cambridge, MA: Academic Press; 2018.
24. Ricci L, Taffoni F, Formica D. On the orientation error of IMU: investigating static and dynamic accuracy targeting human motion. PLoS One 2016;11(9):e0161940 [FREE Full text] [doi: 10.1371/journal.pone.0161940] [Medline: 27612100]
25. Zhang Y, Haghighi PD, Burstein F, Yap LW, Cheng W, Yao L, et al. Electronic skin wearable sensors for detecting lumbar-pelvic movements. Sensors (Basel) 2020 Mar 09;20(5):1510 [FREE Full text] [doi: 10.3390/s20051510] [Medline: 32182928]
26. Papi E, Koh WS, McGregor AH. Wearable technology for spine movement assessment: a systematic review. J Biomech 2017 Dec 07;64:186-197 [FREE Full text] [doi: 10.1016/j.jbiomech.2017.09.037] [Medline: 29102267]
27. Yun WS, Kim H, Ahn JH, Park YB, Park YJ. Individual characteristics of reliable lumbar coupling motions. Eur Spine J 2015 Oct;24(9):1917-1925. [doi: 10.1007/s00586-015-4081-0] [Medline: 26100279]
28. Loh KJ, Lin YA. Smart elastic fabric tape for distributed skin strain, movement, and muscle engagement monitoring. U.S. Utility Patent Application. 2021. URL: https://patentimages.storage.googleapis.com/a8/81/5b/4117f6d1403cc2/US20220087565A1.pdf [accessed 2024-02-05]
29. Loh KJ, Kim J, Lynch JP, Kam NW, Kotov NA. Multifunctional layer-by-layer carbon nanotube–polyelectrolyte thin films for strain and corrosion sensing. Smart Mater Struct 2007 Feb 09;16(2):429-438. [doi: 10.1088/0964-1726/16/2/022]
30. Loh KJ, Lynch JP, Shim BS, Kotov NA. Tailoring piezoresistive sensitivity of multilayer carbon nanotube composite strain sensors. J Intell Mater Syst Struct 2007 Oct 17;19(7):747-764. [doi: 10.1177/1045389x07079872]
31. Loyola BR, La Saponara V, Loh KJ. In situ strain monitoring of fiber-reinforced polymers using embedded piezoresistive nanocomposites. J Mater Sci 2010 Jul 31;45(24):6786-6798. [doi: 10.1007/s10853-010-4775-y]
32. Wang L, Loh KJ. Wearable carbon nanotube-based fabric sensors for monitoring human physiological performance. Smart Mater Struct 2017 Apr 13;26(5):055018. [doi: 10.1088/1361-665x/aa6849]
33. Lin YA, Zhao Y, Wang L, Park Y, Yeh YJ, Chiang WH, et al. Graphene K‐Tape meshes for densely distributed human motion monitoring. Adv Mater Technol 2021 Jan;6(1):2000861. [doi: 10.1002/admt.202000861]
34. Lin YA, schraefel MC, Chiang WH, Loh KJ. Wearable nanocomposite kinesiology tape for distributed muscle engagement monitoring. MRS Advances 2021 Jan 28;6(1):6-13. [doi: 10.1557/s43580-021-00005-4]
35. Lin YA, Mhaskar Y, Silder A, Sessoms PH, Fraser JJ, Loh KJ. Muscle engagement monitoring using self-adhesive elastic nanocomposite fabrics. Sensors (Basel) 2022 Oct 07;22(18):6768 [FREE Full text] [doi: 10.3390/s22186768] [Medline: 36146120]
36. Gombatto SP, Brock T, DeLork A, Jones G, Madden E, Rinere C. Lumbar spine kinematics during walking in people with and people without low back pain. Gait Posture 2015 Oct;42(4):539-544. [doi: 10.1016/j.gaitpost.2015.08.010] [Medline: 26380913]
37. Hodges PW, van den Hoorn W. A vision for the future of wearable sensors in spine care and its challenges: narrative review. J Spine Surg 2022 Mar;8(1):103-116 [FREE Full text] [doi: 10.21037/jss-21-112] [Medline: 35441093]
38. Davis FD. Perceived usefulness, perceived ease of use, and user acceptance of information technology. MIS Q 1989 Sep;13(3):319-340. [doi: 10.2307/249008]
39. Bunniss S, Kelly DR. Research paradigms in medical education research. Med Educ 2010 May;44(4):358-366. [doi: 10.1111/j.1365-2923.2009.03611.x] [Medline: 20444071]

XSL•FO
RenderX

40. O'Brien BC, Harris IB, Beckman TJ, Reed DA, Cook DA. Standards for reporting qualitative research: a synthesis of recommendations. Acad Med 2014 Oct;89(9):1245-1251 [FREE Full text] [doi: 10.1097/ACM.0000000000000388] [Medline: 24979285]
41. Holden RJ, Karsh BT. The technology acceptance model: its past and its future in health care. J Biomed Inform 2010 Mar;43(1):159-172 [FREE Full text] [doi: 10.1016/j.jbi.2009.07.002] [Medline: 19615467]
42. Lange J, Schneider A, Jordi C, Lau M, Disher T. Formative study on the wearability and usability of a large-volume patch injector. Med Devices Evid Res 2021 Nov;Volume 14:363-377. [doi: 10.2147/mder.s337670]
43. Guest G, Namey E, McKenna K. How many focus groups are enough? Building an evidence base for nonprobability sample sizes. Field Methods 2016 Jul 24;29(1):3-22 [FREE Full text] [doi: 10.1177/1525822x16639015]
44. Nowell LS, Norris JM, White DE, Moules NJ. Thematic analysis: striving to meet the trustworthiness criteria. Int J Qual Methods 2017 Oct 02;16(1) [FREE Full text] [doi: 10.1177/1609406917733847]
45. Hamilton AB, Finley EP. Qualitative methods in implementation research: an introduction. Psychiatry Res 2019 Oct;280:112516 [FREE Full text] [doi: 10.1016/j.psychres.2019.112516] [Medline: 31437661]
46. Carter N, Bryant-Lukosius D, DiCenso A, Blythe J, Neville AJ. The use of triangulation in qualitative research. Oncol Nurs Forum 2014 Oct;41(5):545-547. [doi: 10.1188/14.ONF.545-547] [Medline: 25158659]
47. Su WC, Yeh SC, Lee SH, Huang HC. A virtual reality lower-back pain rehabilitation approach: system design and user acceptance analysis. In: Proceedings of the Universal Access in Human-Computer Interaction. Access to Learning, Health and Well-Being. 2015 Presented at: Universal Access in Human-Computer Interaction. Access to Learning, Health and Well-Being; August 2-7, 2015; Los Angeles, CA. [doi: 10.1007/978-3-319-20684-4_37]
48. Lin WY, Chou WC, Tsai TH, Lin CC, Lee MY. Development of a wearable instrumented vest for posture monitoring and system usability verification based on the technology acceptance model. Sensors (Basel) 2016 Dec 17;16(12):2172 [FREE Full text] [doi: 10.3390/s16122172] [Medline: 27999324]
49. Blanchard A, Nguyen SM, Devanne M, Simonnet M, Le Goff-Pronost M, Rémy-Néris O. Technical feasibility of supervision of stretching exercises by a humanoid robot coach for chronic low back pain: the R-COOL randomized trial. Biomed Res Int 2022 Mar 9;2022:5667223-5667210 [FREE Full text] [doi: 10.1155/2022/5667223] [Medline: 35309176]
50. Luciani B, Braghin F, Pedrocchi AL, Gandolla M. Technology acceptance model for exoskeletons for rehabilitation of the upper limbs from therapists' perspectives. Sensors (Basel) 2023 Mar 03;23(3):1721 [FREE Full text] [doi: 10.3390/s23031721] [Medline: 36772758]
51. Sun X, Zhao C, Li H, Yu H, Zhang J, Qiu H, et al. Wearable near-field communication sensors for healthcare: materials, fabrication and application. Micromachines (Basel) 2022 May 17;13(5):784 [FREE Full text] [doi: 10.3390/mi13050784] [Medline: 35630251]
52. Andrýsková A, Lee JH. The guidelines for application of kinesiology tape for prevention and treatment of sports injuries. Healthcare (Basel) 2020 May 26;8(2):144 [FREE Full text] [doi: 10.3390/healthcare8020144] [Medline: 32466467]
53. Cheatham SW, Baker RT, Abdenour TE. Kinesiology tape: a descriptive survey of healthcare professionals in the United States. Int J Sports Phys Ther 2021 Jul 01;16(3):778-796 [FREE Full text] [doi: 10.26603/001c.22136] [Medline: 34123530]
54. Yun-An, L, Noble E, Chin-Hsiung L, Loh KJ. Respiration monitoring using a motion tape chest band and portable wireless sensing node. J Commerce Biotechnol 2022;27(1). [doi: 10.5912/jcb1026]
55. George SZ, Lentz TA, Beneciuk JM, Bhavsar NA, Mundt JM, Boissoneault J. Framework for improving outcome prediction for acute to chronic low back pain transitions. Pain Rep 2020;5(2):e809 [FREE Full text] [doi: 10.1097/PR9.0000000000000809] [Medline: 32440606]
56. Collado-Mateo D, Lavín-Pérez AM, Peñacoba C, Del Coso J, Leyton-Román M, Luque-Casado A, et al. Key factors associated with adherence to physical exercise in patients with chronic diseases and older adults: an umbrella review. Int J Environ Res Public Health 2021 Mar 19;18(4):2023 [FREE Full text] [doi: 10.3390/ijerph18042023] [Medline: 33669679]
57. Palazzo C, Klinger E, Dorner V, Kadri A, Thierry O, Boumenir Y, et al. Barriers to home-based exercise program adherence with chronic low back pain: patient expectations regarding new technologies. Ann Phys Rehabil Med 2016 May;59(2):107-113 [FREE Full text] [doi: 10.1016/j.rehab.2016.01.009] [Medline: 27050664]

Abbreviations

APTA: American Physical Therapy Association
FG: focus group
HIPAA: Health Insurance Portability and Accountability Act
IMU: inertial measurement unit
LBP: low back pain
PT: physical therapy
RQA: rapid qualitative analysis
TAM: Technology Acceptance Model

XSL•FO
RenderX

Edited by A Kushniruk; submitted 08.12.23; peer-reviewed by O Ang, S Oh; comments to author 06.01.24; revised version received 13.01.24; accepted 17.01.24; published 29.02.24.

Please cite as:
Lee A, Dionicio P, Farcas E, Godino J, Patrick K, Wyckoff E, Loh KJ, Gombatto S
Physical Therapists' Acceptance of a Wearable, Fabric-Based Sensor System (Motion Tape) for Use in Clinical Practice: Qualitative Focus Group Study
JMIR Hum Factors 2024;11:e55246
URL: https://humanfactors.jmir.org/2024/1/e55246
doi:10.2196/55246
PMID:38421708

©Audrey Lee, Patricia Dionicio, Emilia Farcas, Job Godino, Kevin Patrick, Elijah Wyckoff, Kenneth J Loh, Sara Gombatto. Originally published in JMIR Human Factors (https://humanfactors.jmir.org), 29.02.2024. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on https://humanfactors.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Physicians' and Patients' Expectations From Digital Agents for Consultations: Interview Study Among Physicians and Patients

Andri Färber[1,2], MSc; Christiane Schwabe[2], PhD; Philipp H Stalder[1], MD; Mateusz Dolata[2], PhD; Gerhard Schwabe[2], Prof Dr

[1]ZHAW School of Management and Law, Zurich University of Applied Sciences, Winterthur, Switzerland

[2]Department of Informatics, University of Zurich, Zurich, Switzerland

Corresponding Author:
Andri Färber, MSc
ZHAW School of Management and Law
Zurich University of Applied Sciences
Theaterstr. 17
Winterthur, 8401
Switzerland
Phone: 41 79 642 08 85
Email: xfbe@zhaw.ch

Abstract

Background: Physicians are currently overwhelmed by administrative tasks and spend very little time in consultations with patients, which hampers health literacy, shared decision-making, and treatment adherence.

Objective: This study aims to examine whether digital agents constructed using fast-evolving generative artificial intelligence, such as ChatGPT, have the potential to improve consultations, adherence to treatment, and health literacy. We interviewed patients and physicians to obtain their opinions about 3 digital agents—a silent digital expert, a communicative digital expert, and a digital companion (DC).

Methods: We conducted in-depth interviews with 25 patients and 22 physicians from a purposeful sample, with the patients having a wide age range and coming from different educational backgrounds and the physicians having different medical specialties. Transcripts of the interviews were deductively coded using MAXQDA (VERBI Software GmbH) and then summarized according to code and interview before being clustered for interpretation.

Results: Statements from patients and physicians were categorized according to three consultation phases: (1) silent and communicative digital experts that are part of the consultation, (2) digital experts that hand over to a DC, and (3) DCs that support patients in the period between consultations. Overall, patients and physicians were open to these forms of digital support but had reservations about all 3 agents.

Conclusions: Ultimately, we derived 9 requirements for designing digital agents to support consultations, treatment adherence, and health literacy based on the literature and our qualitative findings.

(JMIR Hum Factors 2024;11:e49647) doi:10.2196/49647

KEYWORDS

adherence to treatment; digital agents; eHealth; electronic medical records; health literacy; mobile health; mHealth; mobile phone

Introduction

Motivation

Consultations are less productive than what physicians and patients would wish [1,2], which hampers health literacy, shared decision-making, and treatment adherence. The recent rise of generative artificial intelligence (AI), such as ChatGPT, has sparked the interest of digital health developers, as they explore how this technology can improve shared decision-making, physician-patient communication, adherence to treatment, and health literacy. In this study, we sought to discover what physicians and patients expect from digital agents (functional requirements) and how this functionality should be provided (nonfunctional requirements). A user-centric perspective is essential for guiding the development of digital agents because it prepares physicians for changes in their consultation methods and allows patients to understand what the new technology can offer.

XSL•FO
RenderX

Through in-depth interviews (refer to the Methods section), we described 3 digital agents to physicians and patients, analyzed their impressions and expectations (refer to the Results section), and deduced a set of design requirements (refer to the Discussion section). An introduction to the related work and concepts for the 3 different digital agents is provided in the following sections.

Related Work and Concepts

Relevant Medical Concepts

Overall, four medical concepts are essential when supporting medical consultations with digital agents: (1) shared decision-making, (2) physician-patient communication, (3) adherence to treatment, and (4) health literacy.

Consultations involve a participatory process between patients and physicians to reach an agreement regarding treatment goals and their implementation [3,4]. "Shared decision-making" has emerged as the gold standard for this participatory process [5-10] as it strives to reach a mutual agreement about therapy [6,7]. However, a systematic review of shared decision-making regarding clinical decisions found that the humanistic aspects of physician-patient communication were rarely assessed [11]. Good "physician-patient communication" is not only about technique or process but also involves understanding the whole person, finding common ground, and enhancing the patient-physician relationship [4]. In this way, physician-patient communication can have a therapeutic effect and influence health benefits [12].

The therapeutic process continues after the patient has left the consultation [3]. Once at home, it is up to the patient to implement the therapy plan, and the extent to which this occurs is referred to as "adherence to treatment" [13]. Adherence focuses on patients taking responsibility for their treatment and physicians collaborating more with their patients [14,15]. However, despite some progress, adherence to treatment remains insufficient [13,16-18]. First, there is a lack of "health literacy" when following the given instructions. Physicians may explain medical issues and treatment options during consultations, but their time is limited, and they must convey as much information as possible. Second, patients are in a stressful situation, which restricts their ability to absorb and hinders their recall [19-24]. Third, physicians may use medical terminology [25] with the following consequences: patients either do not understand or quickly forget what was discussed [26,27]. Brochures and leaflets are typically used to support health literacy, and modern approaches include video, multimedia, computer-assisted learning, mobile apps, and other web-based aids [28-32].

Digital Agents

Digital agents are computers that undertake tasks previously performed by humans. As such, they function autonomously, react to environmental situations, initiate actions, communicate with humans or machines, and behave intelligently [33]. An increasing volume of digitized data, improved algorithms, and better hardware has vastly enhanced the range of tasks that digital agents can perform. The most noticeable aspect is the recent success of generative AI. Nevertheless, the expanding capabilities of digital agents also raise concerns about AI in general and digital agents in particular [34]. Examples include their potential misuse, how they can be controlled, and whether they exhibit bias [35]. Besides these general concerns, researchers are interested in understanding exactly how digital agents interact with humans. Although humanlike behavior may be helpful in some situations, task performance may be impeded by excessive humanness [36,37] such as in situations where humans prefer a digital agent with a background function. This issue is critical in institutional settings [38], where professionalism is vital.

Discussion about the capabilities of digital agents and their suitability has also reached the medical domain [33,39,40]. Conceptually, the dyadic physician-patient consultation becomes triadic [41-44] if a digital agent is included. The presence of digital agents changes the consultation dynamics [45,46] and alters how patients and physicians behave [41]. Despite such insights, the discussion lacks a clear conceptualization of the digital agent's role in the professional context of physician-patient consultation. Consequently, discussing what physicians and patients expect from digital agents during and between consultations has not been possible.

Current Digital Support for Consultation, Adherence to Treatment, and Health Literacy

Physicians use electronic medical records (EMRs) and encounter patient decision aids (PDAs) during consultations, which provides patients access to their data through patient portals. Patients may also store data in their personal health records (PHRs) and take advantage of mobile health (mHealth) apps between sessions.

EMRs support physicians in documenting medical history, including physical examinations and laboratory results. They are intended to reduce costs, improve patient safety, increase efficiency [47], and safeguard data [48,49]. As EMRs are designed primarily for documentation purposes [50], it is the physician's responsibility to determine how to use them in patient interactions. Proper use of EMRs by trained health care professionals can improve health literacy and adherence to treatment compared with paper-based records [51], for example, if physicians share their EMR screens with patients during consultations [52,53]. However, when used ineptly, physicians lose control of the consultation owing to increased gaze shifts and multitasking, which hinders their medical reasoning [47,54]. In the presence of a computer, preexisting positive and negative communication skills are amplified [55,56].

Encounter PDAs support physician-patient consultations by providing decision-related information and choices [57-61]. Although they tend to be simple in design [61], physicians complain that lack of training and experience and insufficient content and format impede meaningful use of encounter PDAs [57,58]. Another challenge is keeping encounter PDAs updated with the latest information [60].

Patient portals provide patients with access to their data stored in EMRs [62]. In such tethered patient portals, the responsibility for maintaining the data lies with the physician. To be understood by patients, information from EMRs must be

XSL•FO
RenderX

translated [62], and this applies to language, graphs, and other multimedia material.

Unlike patient portals, in electronic PHRs, patients themselves enter and maintain their health data [63]. Although PHRs can accumulate more information than patient portals, quality control and manageability are challenging. There is a consensus that more needs to be done (eg, patients also need to understand what they get from the PHR and need to act on what they understand) to enhance health outcomes or treatment adherence than just providing patients with access to their data [64,65]. Better-informed patients are not necessarily healthier patients [64], but there is (1) value and (2) potential in patient portals and PHRs. First, patients want access to their data to review it again at home, discuss it with their families, and use it as a starting point for further online research [62,64]. Second, there is evidence suggesting that patient portals and PHRs are more effective when they are interactive, when they are combined with other services such as reminders or interactive decision support, and when physicians actively promote their use [62,64].

Digital interventions based on mHealth apps promise to support patients' health literacy and adherence to treatment. In 2017, >300,000 health apps were available in online app stores [66]. Not all are considered effective, convenient, or of high quality [67-69], and many have low success rates and high dropout rates [70-72]. Nevertheless, despite their limitations, mHealth apps appear to support patients effectively in treatment adherence [67,73,74]. If they pass the medical quality requirements, they can even be prescribed in the same manner as medicine [75,76]. Physicians are best placed to assist with their use, but this requires their integration into workflows and EMRs [74,77,78], and the security of patient data must be guaranteed [79].

Digital Agents to Support Consultation, Adherence to Treatment, and Health Literacy

Overview

We conceptualized 3 general roles for digital health agents, which tie together the modern medical concepts and previous studies of digital agents with current digital support for consultation, adherence to treatment, and health literacy. These served as a basis for our empirical study, when introducing our selected physicians and patients to digital agents.

A digital agent can be a "digital expert" that provides the right aids at the right time or offers a second opinion about diagnosis and treatment. It can stay in the background of the consultation as a "silent digital expert" or actively participate in the consultation as a "communicative digital expert." Alternatively, it can be a "digital companion" (DC), which supports the patient between consultations. DCs provide patients with comprehensible information about diagnosis and ongoing treatment.

Silent Digital Expert

This is an extension of EMRs, providing the physician with contextual and real-time advice and additional information. The silent digital expert is designed to free the physician from searching vast information sources and allows more time for face-to-face consultation, thereby improving physician-patient communication [4,12]. For example, the silent digital agent can alert physicians to different diagnoses and drug interactions or offer prompts for further questions. The silent digital agent also supports diagnosis and suggests appropriate treatment in a shared decision-making process [5-10]. It acts as an aid to the physician and is visible and accessible only to the physician, and with patient consent, it can record, transcribe, analyze, and summarize the consultation.

Communicative Digital Expert

As the third party in a triadic consultation, the communicative digital expert offers the same functionality as the silent digital expert. However, it actively participates in the consultation by extending the functionality of EMRs and encounter PDAs through an agency. It may be physically represented as a humanlike robot, smart speaker, or device of any shape. As the third party, the communicative digital expert can be invited to comment about the decision-making process of physicians or patients [5-10] and become active in explaining medical topics, thereby improving health literacy [80-83]. As such, it can be considered as a physician's assistant or patient's advocate, thus improving physician-patient communication [4,12]. For example, it might interrupt the dialogue if a physician is very brief or dominant, thereby providing both parties with further information, diagnosis considerations, and treatment recommendations. It acts in an empathetic, patient-centered manner and is capable of identifying and taking patient preferences into consideration.

Digital Companion

This agent is intended to support patients *between* consultations by extending patient portals and PHRs and combining them with an mHealth app. It relies on data from EMRs and supports patient treatment behavior. Its primary goals are to improve the recall of recommendations and information, promote health literacy [80-83], and support treatment adherence [12-18,84]. DC captures the critical points of the physician-patient consultation, translates them into everyday language, enriches them with multimedia elements (audio, picture, diagram, and video), and makes them conveniently accessible to patients or their families at any time. It also provides the patient with curated additional information and interactively supports their health care education based on individual preferences. Using sensor data from various devices (eg, smartphones, smartwatches, pedometers, and blood glucose monitors) and patient's interaction with DC, adherence to the treatment plan is measured, analyzed, and fed back to the patient (and with the patient's consent, to the physician). DC provides context-specific, adaptive interventions [85-88] based on adherence measurement, individual treatment agreement, and patient preferences. For example, adherence support might include diet recipes, exercise instructions, morale-boosting talks, and so on.

XSL•FO
RenderX

Methods

Research Approach

This study aims to understand what physicians and patients require from digital agents. These requirements should be grounded not only on technical vision but also on current consultation practices, with a focus on problem-solving.

Our research approach was inspired by the practice-oriented approach popular in computer-supported cooperative work (CSCW). CSCW is an interdisciplinary field of research involving, among others, computer science, psychology, and sociology, to analyze the potential and the shortcomings of digital assistance in consultations [89-91]. CSCW mainly uses qualitative methods and focuses on how human collaboration can be supported by technical means [89,92]. As these means must be applied within a professional context, this also involves studying work practices from the perspective of those involved [93,94].

Our study embraced this tradition by following an exploratory paradigm, striving for deep, contextualized insights [95,96]. We conducted an interview-based qualitative study with 47 participants—22 (47%) physicians and 25 (53%) patients. Our analysis combined bottom-up thematic analysis and interpretive research, allowing for both broad coverage and deep insight.

Overall, the chosen methodological approach respected the need to understand patients' and physicians' perspectives regarding their work practices and the potential use of technologies. We addressed variation and triangulation, whereby multiple researchers conducted the interviews with different patients and physicians. We ensured audit throughout the process by mutual control among researchers and by assigning a quality manager role to one of the authors. The first author was directly engaged in data collection during a preliminary study [97] and guided data collection during this study to ensure adequate engagement in data collection activities. In summary, the study used various strategies to ensure the reliability and validity of the presented results [98] and followed the COREQ (Consolidated Criteria for Reporting Qualitative Research) guidelines for reporting qualitative research [99].

Ethical Considerations

The Ethics Committee of the Zurich canton confirmed that this study was not subject to the Swiss Human Research Act (Business Administration System for Ethics Committees [BASEC]–Nr Req-2018-00847). Nevertheless, written informed consent was obtained from all participants before their interviews according to the World Medical Association Declaration of Helsinki [100].

Sampling and Recruitment

Exploratory studies require a variety of opinions, but they do not seek to be representative. To ensure variety, we interviewed both physicians and patients. We also relied on purposive sampling using a maximum variation strategy [101], which allowed us to search for a broad range of physicians and patients. Given that 5 interviewers acquired the patients and physicians independently, we can assume the coverage to be better than that of strategies involving sampling through a single researcher. Table 1 shows the demographic characteristics of the study participants.

Table 1. Demographic data of the interviewed physicians and patients.

Characteristics	Physicians (n=22)	Patients (n=25)
Sex, n (%)		
Male	12 (55)	14 (56)
Female	10 (45)	11 (44)
Age (y), mean (SD; range)	50 (14; 25-66)	46 (19; 20-86)

Of the 22 physicians, 13 (59%) are active in primary care, and the others work in hospitals; 11 (50%) are general practitioners or specialize in internal medicine. Other specializations include pediatrics, gynecology, radio-oncology, and dentistry. The educational background of the 25 patients ranged from unskilled workers to professionals and academics. The patients presented a broad spectrum of conditions, including diabetes, multiple sclerosis, heart conditions, tick-borne encephalitis, and epilepsy.

We conducted 46 in-depth interviews that resulted in audio recordings with 32 hours of interview time, amounting to an average length of 42 minutes and 46 seconds (SD 13 min and 47 s). Of the 46 interviews, 45 (98%) were conducted with 1 interviewee per session, and 1 (2%) involved 2 respondents. The sample size assured data saturation—the topics emerging in the interviews began to overlap after about 18 to 20 interviews for each group [102]. Consistent with the practice for purposive sampling and maximum variation [101], we used various channels to establish the initial contact with the interviewees (email, face-to-face, and telephone). After confirming the time and date for a potential interview and giving their consent, no one dropped out of the study.

Data Collection

In total, 5 researchers conducted in-depth interviews based on the respective interview guides—separate guides for patients and physicians [96]. The interview guides were developed based on the literature about physician-patient communication; adherence to treatment; existing solutions in the field of medical informatics; and the authors' own experiences in the medical domain, including their research background. The overall structure of the interviews was informed by CSCW practice-oriented studies [93,94]. The interview guides were pretested in a preliminary study (with 11 health care professionals and 7 patients) published elsewhere [97]. Interviews for this study were conducted between January 2019 and May 2019, with patient interviews being conducted mostly in their homes and health care professional interviews in their

XSL•FO
RenderX

professional setting. Before the interviews, all researchers underwent interview training sessions to ensure that they had the same understanding of the questions and knew how to conduct the interview. The interviews were structured around 3 areas: current situation or practice (format of and preparation for a consultation), future developments (expectations from and attitudes toward digital health care), and closure (other points that were not already covered).

When discussing about digital developments, we suggested potential ideas because users often lack the necessary imagination when asked about future products or services [103]. Nevertheless, when prompted, many users can express helpful, subjective opinions about specific ideas [103]. Therefore, in the spirit of design thinking [104], we exposed the users to key design ideas by describing the digital experts and DC and asking for their perceptions, expectations, and preferences regarding digital agent support. As is typical in design thinking, the discussion focused on the desirability of critical capabilities but did not include a detailed discussion about feasibility.

Data Analysis

All the interviews were audio recorded and transcribed. The analysis combined deductive thematic research and interpretive research, allowing for broad coverage and deep insight simultaneously. During the top-down analysis, the transcripts were coded according to a codebook derived inductively from a small preliminary study [97]. A professor of nursing science cross-checked the codebook. Again, all researchers attended a training session to ensure that they had the same understanding of the codebook. All interviews were then deductively coded using MAXQDA (VERBI Software GmbH) [105]. The designated quality manager conducted quality assurance activities by controlling all code assignments and correcting them to ensure a consistent basis for analysis. We achieved thematic saturation—all themes from the specified coding schema appeared in the data with high frequency (the most frequent code was assigned 274 times and the least frequent was assigned 25 times; overall, we had 1954 assignments across all codes) [102]. Finally, all interviews were summarized by code; for each theme, we obtained a summary of participant opinions related to the code. These summaries formed the basis for further analysis, and the results were then used for interpretation.

To interpret the data, we organized 2 interpretation workshops involving the authors. The workshops aimed to establish a shared and consistent understanding of the most essential insights between the authors. The interpretive process involved iterative restructuring of the summaries along various dimensions, with 2 dimensions emerging as crucial for forming a consistent data view. First, we differentiated the problems, current practices that emerged to mitigate those problems, and potential technological solutions to address the problems that occurred during the interviews. Second, we observed that the issues aligned with the phases of a patient's journey: (1) consultation, (2) "transition" between consultations and period between consultations, and (3) actual period between consultations. These differentiations provided the framework for reporting our results, and the proposed structure covered all the challenges and problems identified during coding.

In our presentation of the results, we refer to the frequency of specific challenges because, after identifying the framework and distributing the significant challenges for each element in the framework, we returned to the coded data to classify the coded passages. In the following section, we have presented the quantified data about the frequency of passages pertaining to the challenges. However, it is important to clarify that we do not assert the representativeness of these figures, as the analyzed population was not chosen to be representative of the broader population. Instead, the numbers ensured the thematic saturation mentioned previously.

Results

Through analysis, we categorized the results into 3 steps in the patient journey: first, the consultation; then, incorporating information from the consultation into their lives; and finally, the time between consultations.

Problems and Agent-Based Solutions During a Consultation

During consultation, the main challenge, according to physicians and patients, is conveying complex information in minimal time to laypeople with various backgrounds, expectations, and abilities while building or maintaining a relationship of trust. Table 2 summarizes the problems voiced by physicians and patients, current practices (as presented by the interview partners), and envisioned solutions offered by the 2 different versions of digital experts.

XSL•FO
RenderX

Table 2. Problems and solutions suggested during a consultation, along with the number of mentions in interviews.

Problems during the consultation	Current practices	Solutions offered by the silent digital experts	Solutions offered by communicative digital experts
Time pressures (physicians: 5/22, 23%; patients: 3/25, 12%)	__[a]	• Physicians can concentrate on a thorough and engaging consultation using digital situational information	• Physicians can concentrate on a thorough and engaging consultation using digital situational information

Medical information is complex (physicians: 9/22, 41%; patients: 18/25, 72%)	• Physicians use graphics, visualizations, videos, and 3D models from brochures, books, and online sources (physicians: 14/22, 64%; patients: 7/25, 28%)
	• Physicians draw illustrations themselves (physicians: 13/22, 59%; patients: 4/25, 16%)
		• The digital expert provides physicians with the following:
		  • The right visual aid at the right time
		  • Graphic templates or blank drawing areas that they can use for their own drawings
			• The digital expert suggests text, images, audios, and videos tailored to individual patient needs

Not all patients respond to medical advice and information in the same manner (physicians: 12/22, 55%; patients: no matching question)	• Most physicians try to approach patients individually by adapting their language to a patient's educational background or medical knowledge (physicians: 4/22, 18%; patients: no matching question)	• The digital expert provides visual aids tailored to a patient's educational background, numerical ability, or language skills	• The digital expert intervenes if it determines (eg, through sentiment analysis [106]) that a patient does not understand the physician
Patients expect more transparency and control over the treatment process (physicians: 2/22, 9%; patients: 21/25, 84%)	• Many patients engage in conversations with physicians and take responsibility for their treatment (patient: 3/25, 12%), and physicians try to support this (physicians: 4/22, 18%; patients: 5/25, 20%)	—	• The digital expert intervenes when physicians do not give their patients enough time to talk, and it can empower patients to take more control
Some patients do not agree with the proposed treatment plan (physicians: 4/22, 18%; patients: 6/25, 24%)	• Physicians respond with more intensive explanations (physicians: 8/22, 36%) • Physicians protect themselves by documenting the conversation • Physicians do not enforce treatment	• The digital expert offers arguments, statistics, and figures to support the physician's point of view	• The digital expert advocates for the patient (by putting the physician's thoughts or guidelines into perspective) or for the physician (by supporting the physician's thoughts or guidelines)
The computer distracts the physician and interrupts communication, and use of computer amplifies inferior communication skills	—	• The user interface of the digital expert is designed to be self-explanatory and user-friendly • Instead of the physician, the digital expert searches for information and offers context-related content	• The digital expert supports the physician and the patient, for example, through active listening • It will only interfere by assisting an already impaired conversation

[a]Nothing mentioned in the interviews.

Regarding current practices, patients and physicians report that there is very little time for a thorough and engaging conversation:

> *I just felt like I was being processed. Quick assessment with the question: What's the problem? And I felt that I couldn't even say what I had because it was already clear to the physician. After a quarter of an hour, I was out of there again, and I was no wiser.* [Male patient; aged 60 years; D07]

> *I frequently make lifestyle recommendations. Costs time too, by the way, cannot be done in a 20-minute consultation that's just long enough for issuing a prescription.* [Male general practitioner; aged 64 years; hospital; ST09]

Most physicians in this sample practice shared decision-making. Some use the explicit term during the interview, whereas others simply implement shared decision-making without labeling it as such:

> *Then I say, we could try pharmacy, we could try herbs, we could try acupuncture or this or that. I'll let the patient have a say. Because then the patient's adherence is also much better.* [Female general practitioner; aged 65 years; medical office; MA10]

All interviewed patients favored a silent digital expert as an aid to the physician; they did not object to physicians using online sources to obtain additional information during a consultation:

> *I don't like having a doctor who introduces him- or herself as "I am the all-knowing one." For me, that*

XSL•FO
RenderX

> *tends to inspire confidence when a physician says: I don't know, I have to work with the exclusion procedure.* [Male patient; aged 74 years; F01]

However, patients expect uninterrupted attention, which requires a sufficiently high level of expertise by the physician in using the computer:

> *He kept asking and reading to me while he was writing and asking me if that was correct. This was great for me because then I knew what he was writing.* [Female patient; aged 52 years; S10]

Most physicians in this study would welcome a silent digital expert to facilitate multitasking, and some already use drug interaction assistants, risk or score calculators:

> *You can't read through the books in the evening. That would mean an insane amount of time or such a head. That's why these are important tools, I think for rare conditions it's certainly a good idea.* [Female gynecologist; aged 35 years; hospital; MA02]

However, the benefits of a digital expert are assessed differently by those in different medical disciplines. A physician was concerned about the transfer of responsibility to the digital expert, whereas another physician worried about a decline in interprofessional communication. A young physician was concerned that this would cause them to acquire very little experience and self-confidence:

> *You rely too firmly on that afterward. Then you believe too firmly in that. Then it takes over your task, so to speak.* [Female dentist; aged 29 years; dental surgery; MA03]

Most patients in this sample view communicative digital experts positively. Those against them are concerned that they might be disruptive or could be manipulated by the physician:

> *I do not know what the physician can enter there, and then it is clear that the computer represents the opinion of the physician.* [Female patient; aged 51 years; S07]

The opinions of those in favor of it differ. Some consider a communicative digital expert as helping less skillful physicians and others consider it as helping competent physicians. Some would like a digital expert to be a physician's assistant, whereas others consider it as a patient's advocate:

> *As a patient, you are always subordinate to the physician, in that sense. I don't think it's a bad thing when someone else is on my side.* [Female patient; aged 28 years; S06]

Approximately two-thirds of the interviewed physicians reject the communicative digital expert. For them, credibility, decision-making authority, and their patients' trust are at stake. Some consider empathy between the physician and patient as essential for patient adherence to treatment and, therefore, do not believe that a digital expert can help. A physician found communicative digital experts annoying but assumed that physicians and patients would get used to them over time:

> *In principle, I say, there is still an interpersonal level that artificial intelligence cannot comprehend.* [Female general practitioner; aged 48 years; medical office; MA08]

Problems and Agent-Based Solutions for Transitioning From Consultations to the Period Between Consultations

Problems during the consultation may also hinder treatment because poor consultations can impair health literacy and adherence to treatment. Table 3 provides an overview of the voiced consultation issues that affect the time between consultations and the envisioned solutions offered through an interaction of the digital expert and DC.

Table 3. Problems and envisioned solutions for transitioning from consultations to the period between consultations, along with the number of mentions in the interviews.

Problems resulting from the consultation	Current practices	Solutions offered by the digital experts connecting to the digital companion
Patients cannot remember everything that the physician says (physicians: 0/22, 0%; patients: 10/25, 40%)	• Patients do the following: • Bring companions to the consultation • Consult brochures or online sources (patients: 2/25, 8%) • Use reminders on smartphones (patients: 2/25, 8%) • Take notes (patients: 6/25, 24%) • Physicians do the following: • Repeat (physicians: 2/22, 9%) • Use active listening techniques	The digital expert records, transcribes, and summarizes the conversation for the patient (quality assurance)
Identifying and introducing clinically relevant mHealth[a] apps is time consuming and difficult	• Patients search for apps themselves, but use dropout rates are high	The digital expert suggests quality-assured mHealth apps or equivalent features of the digital companion

[a]mHealth: mobile health.

Most physicians in this sample see potential in automated recording and transcription. A physician hoped that digital experts would give them more time to communicate with patients. However, physicians doubt whether a computer can separate relevant statements from irrelevant ones and produce relevant summaries. Some physicians stress that the notes they

XSL•FO
RenderX

make for themselves about the case cannot be directly shared with the patient but need to be translated. Others insist on control over the information that is shared with patients:

> *Therefore, the software must either be able to guarantee this or otherwise it is legally difficult to prove that the patient has been informed correctly.* [Male radio-oncologist; aged 35 years; hospital; MA01]

Besides technical difficulties, the interviewed physicians see another reason to avoid automatic summaries—subjective perceptions are often only discussed verbally or communicated via telephone owing to fear of litigation:

> *Certain things, incidents and so on, or special experiences or special stories that are told that could have legal relevance. I don't list them in the computer.* [Male general practitioner; aged 62 years; medical office; ST02]

Another physician takes precisely the opposite position. They would appreciate transcripts of complex consultations in which, for example, discussions about child protection or off-label prescriptions of medication are involved. A physician did not believe that a consultation's significant first and last seconds would be transcribed with the necessary weighting.

Patients also have different opinions about digital experts. Only a few patients in this study raised data protection concerns regarding the consultation transcripts and other information recorded during the consultation. Some patients indicated that they would benefit from this evidence of what was said in the event of disagreement or malpractice. A patient was worried about a decline in care because physicians were afraid of malpractice lawsuits:

> *I tend to think I get worse treatment because most physicians have way too much fear of someone coming in afterward and saying, "I'm going to sue you – you told me something wrong."* [Male patient; aged 61 years; S02]

Problems and Agent-Based Solutions for the Period Between Consultations

The consultation cannot cover all the questions and issues arising between consultation appointments, and patients must rely on their own judgment or a tool that assists them during this period. Table 4 presents the problems that arise between consultations that lead to poor adherence and the solutions offered by DC.

XSL•FO
RenderX

Table 4. Problems and envisioned solutions for the period between consultations, along with the number of mentions in the interviews.

Problems arising between consultations	Current practices	Solutions offered by DC[a]
• Patients lack information because of the following: • Insufficient time for explanations during the consultation (physicians: 5/22, 23%; patients: 3/25, 12%) • Poor recall of the consultation (physicians: 3/22, 14%; patients: 11/25, 44%) • More questions arising later (physicians: 0/22, 0%; patients: 11/25, 44%)	• Patients do the following: • Use online sources, but they are skeptical, and some distrust online forums in particular (physicians: 6/22, 27%; patients: 6/25, 24%) • Read brochures (patients: 3/25, 12%), attend public lectures, or even attend anatomy courses • Physicians provide brochures to guide patients away from online self-diagnosis (physicians: 2/22, 9%).	DC provides curated content and web links tailored to the patient's diagnosis. This reduces misinformation and false self-diagnosis. In addition, it fosters more trust in health care information.
• Patients lack clear instructions and specific information but instead experience information overload (physicians: 0/22, 0%; patients: 8/25, 32%)	• Physicians provide paper-based instructions regarding medication, exercises, and lifestyle changes (physicians: 3/22, 14%; patients: 8/25, 32%)	DC tailors content to patient preferences, contexts, and specific circumstances. This includes content presentation in different formats (simple or sophisticated text, images, audios, and videos).
• Patients are on their own between consultations (physicians: 3/22, 14%; patients: 1/25, 4%)	• Patients report little interaction with their physicians between consultations • Some use email but only sparingly (physicians: 2/22, 9%; patients: 7/25, 28%)	DC provides low-barrier access to the physician between consultations. A chatbot covers part of the conversation to protect physicians from huge workload.
• Patients are overchallenged when taking their medication (physicians: 2/22, 9%; patients: 8/25, 32%)	• Medication apps can support complicated medication regimes (physicians: 1/22, 5%; patients: 2/25, 8%)	DC supports adherence by providing the patient with individualized interventions that consider patient preferences, contexts, and specific circumstances.
• Treatment success or failure goes unnoticed (physicians: 4/22, 18%; patients: 0/25, 0%)	• Physicians ask patients to maintain diaries or journals, mostly paper based (physicians: 8/22, 36%; patients: 3/25, 12%)	DC offers easy-to-maintain diaries and journals, including data captured from digital devices (eg, wearables). The collected data can be shared with physicians (with the patient's consent).
• Measuring adherence is difficult	• Adherence is rarely measured, and often, it is only based on the purchase of medicines (physicians: 1/22, 5%; patients: no corresponding question)	DC offers adherence measurements in an easy-to-understand format.

[a]DC: digital companion.

Most patients in this study would welcome a DC; however, a few are skeptical or undecided. Patients are open to using electronic tools and online services regarding current practices. However, this is not always helpful to physicians:

> *People practically come with a diagnosis, and after that, we first have to come back to the symptoms. And I have to say, "hey, we have to start all over again."* [Male general practitioner; aged 66 years; medical office; ST01]

Many physicians who were interviewed could see the potential of a DC. Some hoped this would improve adherence to medical advice, whereas a physician saw a significant benefit in making the DC genuinely personalized and tailored to an individual patient's needs. Regarding monitoring patient behavior between consultations, less than one-third of the physicians reported adherence measurement (which is usually based on the purchase of medications):

> *That's why I'm very happy when the patients order medication from us because then I can see on the computer when they have picked up their medication. I don't see that when they buy medicines from the pharmacy.* [Female general practitioner; aged 48 years; medical office; MA08]

Most physicians in this sample are open to receiving and interpreting monitoring data from patients and their mobile devices. However, they have the following reservations. First, there is an unmanageable number of mobile apps. Second, they fear data overload and being forced to respond to monitoring results, which requires additional time that physicians do not have. Third, physicians see a risk that such monitoring will negatively influence patient behavior. A physician raised the possibility that neurosis could result from constant introspection. Another concern was that patients would abdicate responsibility for their condition by transmitting data and threshold violations. Despite these concerns, confronting patients regarding their threshold violations encourages them to reflect on their condition and possible lifestyle changes. Therefore, patients can become "experts" on their condition:

XSL•FO
RenderX

Because that is certainly one aspect when patients think about it: Why did my sugar do that now? That's the most instructive. And the goal is that they become the "expert" and I coach them. [Female general practitioner; aged 39 years; hospital; ST08]

Discussion

Overview

Problems in physician-patient interaction that ultimately hamper treatment adherence can be classified into 3 categories: problems regarding the consultation itself, problems from the consultation but appearing between consultations, and new problems arising between consultations. These problems overlap and, therefore, need to be addressed using integrated support systems. On the basis of the scenario, a support system consisting of digital agents assisting in the consultation and a companion for the periods between consultations is proposed. To qualify for the task, these agents need to meet the expectations of physicians and patients and improve health outcomes. In the following sections, we discuss design recommendations for the 3 digital agents that are active in the consultation and act as the patient's companion between consultations.

Requirements for Digital Experts During the Consultation

Digital experts reveal their capabilities during the consultation by integrating and extending the functionalities of EMRs and encounter PDAs with the characteristics of digital agents [33]. These include autonomous and intelligent behavior, reactions to environmental situations, and communication with humans or machines.

The Digital Agent Should Make Its Role in the Triadic Consultation Transparent

Our interviews asked for opinions about including medically skilled digital agents as part of a physician's EMR [45,46]. These can facilitate conversations between physicians and patients or offer second opinions regarding diagnosis and treatment. In such cases, the digital agent functions as an additional physician. Although most patients would welcome this triadic consultation, some fear that physicians could manipulate their DCs. These reservations arise from an understanding that digital agents could adopt the role of a second physician and a trusted family member, spouse, or friend [41,42]. Such roles include informational or emotional support (eg, taking notes, ensuring understanding, and reassuring patients) [42]. Accordingly, the role of a digital agent in consultation must be clearly defined and transparent to patients. Further studies might explore what patients require to trust and benefit most from these digital agents in the role of a second physician, family member, spouse, or friend.

The Digital Agent Should Encourage Trust and Support the Physician-Patient Relationship While Safeguarding the Physician's Credibility

The literature and interviews with physicians and patients agree on the importance of trust and good relationships between physicians and patients in a medical setting [4,12]. Although traditional health IT (eg, EMRs and encounter PDAs) does not seem to interfere with patient-physician relationships [53], the situation changes when digital agents act as medical experts or DCs during a consultation. Most interviewed patients like the idea of a digital agent and do not think it will harm the physician-patient relationship. At the same time, many physicians have an opposing view, fearing loss of credibility and decision-making authority. Therefore, a challenge for DC is to foster trust and support, rather than undermine, the relationship between physicians and patients. Such digital agents must support patients but not unduly contradict physicians or disrupt the natural flow of conversation. This means that digital agents must recognize whether a piece of medical advice will strengthen or damage the relationship.

The Digital Agent Should Help Physicians to Focus on the Patient During the Consultation

The interviewed patients expect their physicians' full attention even when interacting with a computer. In a traditional practice setting, computer screens create a barrier between patients and physicians and can be a serious distraction [47,54]. However, digital agents act independently or are triggered by voice control to provide information or document the conversation, requiring less attention from the physician. The form of digital agents integrated into the conversation can range from shared screens or smart speakers to humanlike robots. Technological advances have brought such user interfaces and digital agents more close to reality. Further studies should indicate what patients and physicians are most likely to accept.

The Digital Agent Should Support Physicians by Taking Over Administrative Duties

Administrative duties prevents physicians from doing what they were trained to do (at considerable expense) and reduces their job satisfaction. The time pressure resulting from these administrative duties is a well-known problem that affects patient health outcomes [1,2,12]. This issue surfaced in the interviews with physicians and patients who were dissatisfied with their treatment. Therefore, a significant role for digital experts is to relieve physicians from as many administrative duties as possible. However, it is essential for physicians that their medical reasoning is considered as something more than mere administration. Recording, transcribing, and summarizing the conversation is necessary, but it is not the whole story. Digital experts should support medical reasoning of physicians and ask for it if not already done, rather than impeding it.

Requirements for Handover From Digital Experts to DCs

To ensure a seamless patient experience, information collected and discussed during the consultation must be passed from the digital experts supporting the consultation to a patient's DC.

The Digital Agent Should Tailor Information and Patient Education to Individual Patient Needs and Preferences

In supporting consultation, digital experts could, for example, provide appropriate information at the appropriate time. After consultation, DCs could continue patient education between consultations, which is tailored to their information needs and

XSL•FO
RenderX

preferences. This can give physicians extra time during consultations [1,2] and assist patients in recalling recommendations and information [19,23,24]. In contrast to reading widely circulated brochures, leaflets, and generalized online sources [28,29,31,32,107], patients receive personalized information matching their specific circumstances and treatment plans. This saves time by reducing the need to guide patients away from potentially incorrect self-diagnosis [30].

Our interviews indicated that physicians effectively tailor information to their patients' needs and backgrounds. Therefore, digital agents in the form of digital experts and companions must keep up with or even outperform physicians to add value. To achieve this, digital experts should either be able to draw on predefined patient profiles or interpret and assess patient preferences and backgrounds correctly. Physicians understandably insist on maintaining overall control as they are liable for the information they give their patients. A suboptimal solution would require physicians to verify the information they provide patients via the DC. In contrast, a better solution would ensure (in a trusted manner) that the information offered was consistent with the physician's directions.

Requirements for the DC in the Period Between Consultations

DCs support patients as digital agents between consultations by integrating and extending the functionalities of patient portals, PHRs, and mHealth apps.

The Digital Agent Should Offer Adaptive Interventions for Behavior Change

In conventional lifestyle change treatment, adaptive interventions are standard, and physicians and patients adapt and agree about the treatment every few weeks or months, ideally in a shared decision-making process [3,4,6,7,9]. However, adjustment cycles are dependent on consultation cycles, and in the meantime, patients may treat themselves incorrectly or discontinue a treatment owing to a lack of corrective measures. Here, digital agents in the form of DCs can shorten the cycle considerably. Depending on a patient's mood, context, experience, and feedback, the DC can adjust the treatment within days, hours, minutes, or even seconds [85,86]. In our interviews, patients welcomed the idea of such functional flexibility. However, the challenge for the digital agent is to offer adaptive interventions that align with the respective physician's recommendations, comply with medical device regulations, and fulfill safety and performance requirements. Further studies must demonstrate that this type of adaptive intervention will improve treatment adherence.

The Digital Agent Should Measure and Monitor Patients' Adherence to Treatment and Provide Physicians With Easy-to-Read and Easy-to-Interpret Summaries

Measuring patients' adherence to treatment is a prerequisite for adaptive interventions [13]. Our interviews indicate scope for improvement regarding the measurement of treatment adherence—particularly for exercise and lifestyle changes. DCs are well suited to measure adherence based on objective data from sensors and subjective data such as chatbot conversations with patients. The interviewed physicians indicated that they would accept patient behavior monitoring if DCs aggregated the monitoring results and communicated them directly to EMRs. The literature also calls for this type of workflow integration [62,74,77,78]. However, the DC must be able to recognize red-flag situations and respond appropriately because the responsibility and workload of constantly monitoring the results cannot solely rely on physicians.

Further studies are needed to determine how patients respond to behavioral monitoring. The interviewed physicians anticipate positive effects, such as patients becoming "experts" on their condition, and adverse effects, such as patients relinquishing responsibility for their actions. Therefore, digital agents must monitor patients in a supportive manner and report the results in a form that assists rather than overloads the physician.

The Digital Agent Should React to Feedback and Questions From Patients in the Period Between Consultations

The more sophisticated the DC's communication and interaction skills are, the greater the expectation patients have for them to react appropriately. It is insufficient to simply give patients access to information through patient portals or PHRs [62,63] or have chatbots handling patient questions and feedback. In certain circumstances, patients still wish to talk to their human physician. In such cases, a triage mechanism might involve physicians only when necessary. However, the associated liability issues affecting the physicians (eg, in the case of suicidal intent) must be resolved.

Requirements for the Integration of Digital Experts and DCs

Only the integration of digital experts and DCs can unlock the full potential of these agents to support the entire consultation process for the mutual benefit of patients and physicians.

The Digital Agent Should Integrate Consultation Support (Digital Experts) and Patient Apps (DCs)

Integrating digital experts and DCs closes the loop from one consultation to the next and synergistically increases the benefits of both agents [108]. From a digital expert to a DC, personalized information about the diagnosis and treatment is transmitted immediately at the end of the consultation. This avoids media discontinuity, overcomes the problem of poor recall of recommendations or information, and allows patients to implement correct therapy immediately. Some of this functionality is already part of patient portals or PHRs [62,63]. However, making this information available in an mHealth app supported by digital agents allows for better interactivity, adherence support, and measurement. As access to information alone has not proven to be effective [64,65], the mHealth approach promises greater effectiveness. Adherence measurements are fed from the DC to the digital expert based on sensor data and patient-reported outcome measures (eg, diary entries and chatbot threads). This allows physicians to prepare for the next consultation and saves time because patients do not have to report verbally what they have already entered into the app. The interviewed physicians and patients welcomed this

XSL•FO
RenderX

focus and time-saving measure, and the literature also calls for workflow integration along these lines [77,109-111].

Limitations

We derived the requirements for the design of digital agents to support consultation, adherence to treatment, and health literacy solely based on the statements obtained from our in-depth interviews with patients and physicians. Therefore, the 9 resulting requirements cannot be described as exhaustive. In particular, many necessary nonfunctional requirements are still lacking.

Furthermore, this study was conducted in Switzerland, which has one of the most expensive health care systems in the world. According to participating physicians, the standard consultation time is 20 minutes, which is significantly longer than that in many other countries. The responses from patients and physicians in other places and cultures might differ considerably. Further limitations may have arisen from the nature of a qualitative study based on a purposive sample. Although such a study results in a broad picture and deep insights, it may not be representative, not even for Switzerland. In addition, it is impossible to quantify the importance of the issues, suggested solutions, participant feedback, or the derived design requirements. For such purposes, surveys based on the insights obtained from this study are better suited. In addition, we cannot draw any conclusions related to specific user groups or medical disciplines. The fact that interview partners from very diverse backgrounds made similar observations and judgments indicates that our findings could be applied to various disciplines and user groups.

Conclusions and Future Studies

With the introduction of generative AI such as ChatGPT, the time for digital agents to support consultation, adherence to treatment, and health literacy may have arrived. There is enormous potential for patients and physicians to benefit from this new technology. Through in-depth interviews, both parties revealed their opinions about a silent and a communicative digital expert to support consultation and a DC to accompany patients between consultations. Their responses are synthesized into the following 9 requirements for the design of digital agents to support consultations.

The digital agent should do the following:

1. Make its role in the triadic consultation transparent
2. Encourage trust and support the physician-patient relationship while safeguarding physician credibility
3. Help physicians to focus on the patient during the consultation
4. Support physicians by taking over administrative duties
5. Tailor information and patient education to individual patient needs and preferences
6. Offer adaptive interventions for behavior change
7. Measure and monitor patient adherence to treatment and provide physicians with easy-to-read and easy-to-interpret summaries
8. React to feedback and questions from patients in the period between consultations
9. Integrate consultation support (digital experts) and patient apps (DCs).

Some recommendations for future studies were also offered in Requirements for Digital Experts During the Consultation section and Requirements for the DC Between Consultations section in the Discussion section. In addition, we suggest the following:

1. Obtain a complete set of requirements for the design of digital agents for consultation; a full requirement engineering approach would need to be followed and explored in the field. This would include an analysis of the technical feasibility and economic viability [104] of the system, with the results of this study serving as a starting point.
2. Depending on where the digital agents are to be deployed, this study could be replicated with local patients and physicians.

Acknowledgments

The authors thank Andrea Koppitz Prof Dr, Professor of Nursing, for reviewing the codebook and assisting them with their profound health care knowledge. Furthermore, the authors would like to thank Dominik Spescha and Daniel Degenhardt for their support in conducting, transcribing, and analyzing the interviews.

Authors' Contributions

AF was involved in conceptualization, methodology, validation, investigation, data curation, writing the original draft, review and editing, visualization, and project administration. CS was involved in writing the original draft and review and editing. PHS contributed to the investigation. MD was involved in conceptualization and review and editing. GS contributed to conceptualization; reviewing, editing, and rewriting some sections; and supervision.

Conflicts of Interest

None declared.

References

1. Costello KL. Impact of patient-provider communication on online health information behaviors in chronic illness. Proc Assoc Inf Sci Technol 2016 Dec 27;53(1):1-10. [doi: 10.1002/pra2.2016.14505301060]

XSL•FO
RenderX

2. Golder L, Jans C, Tschöpe S, Venetz A, Herzog N. Verändertes arbeitsumfeld und einstellung zu neuen finanzierungsmodellen: auswirkungen leistungsorientierung im gesundheitswesen erkennbar. gfs.bern. 2018 Jan. URL: https://www.fmh.ch/files/pdf20/2018_02_15_Begleitforschung_FMH_Langversion_SB_final.pdf [accessed 2024-02-27]
3. Kissling B, Ryser P. Die Ärztliche Konsultation Systemisch-Lösungsorientiert. Göttingen, Germany: Vandenhoeck & Ruprecht GmbH & Co KG; 2019.
4. Stewart M, Brown JB, Weston W, McWhinney IR, McWilliam CL, Freeman T. Patient-Centered Medicine: Transforming the Clinical Method, 3rd Edition. Boca Raton, FL: CRC Press; 2013.
5. Gafni A, Charles C, Whelan T. The physician-patient encounter: the physician as a perfect agent for the patient versus the informed treatment decision-making model. Soc Sci Med 1998 Aug;47(3):347-354. [doi: 10.1016/s0277-9536(98)00091-4] [Medline: 9681904]
6. Elwyn G, Frosch D, Thomson R, Joseph-Williams N, Lloyd A, Kinnersley P, et al. Shared decision making: a model for clinical practice. J Gen Intern Med 2012 Oct;27(10):1361-1367 [FREE Full text] [doi: 10.1007/s11606-012-2077-6] [Medline: 22618581]
7. Stiggelbout AM, Pieterse AH, De Haes JC. Shared decision making: concepts, evidence, and practice. Patient Educ Couns 2015 Oct;98(10):1172-1179. [doi: 10.1016/j.pec.2015.06.022] [Medline: 26215573]
8. Shay LA, Lafata JE. Where is the evidence? A systematic review of shared decision making and patient outcomes. Med Decis Making 2015 Jan;35(1):114-131 [FREE Full text] [doi: 10.1177/0272989X14551638] [Medline: 25351843]
9. Bomhof-Roordink H, Gärtner FR, Stiggelbout AM, Pieterse AH. Key components of shared decision making models: a systematic review. BMJ Open 2019 Dec 17;9(12):e031763 [FREE Full text] [doi: 10.1136/bmjopen-2019-031763] [Medline: 31852700]
10. Makoul G, Clayman ML. An integrative model of shared decision making in medical encounters. Patient Educ Couns 2006 Mar;60(3):301-312. [doi: 10.1016/j.pec.2005.06.010] [Medline: 16051459]
11. Kunneman M, Gionfriddo MR, Toloza FJ, Gärtner FR, Spencer-Bonilla G, Hargraves IG, et al. Humanistic communication in the evaluation of shared decision making: a systematic review. Patient Educ Couns 2019 Mar;102(3):452-466. [doi: 10.1016/j.pec.2018.11.003] [Medline: 30458971]
12. Street RLJ, Makoul G, Arora NK, Epstein RM. How does communication heal? Pathways linking clinician-patient communication to health outcomes. Patient Educ Couns 2009 Mar;74(3):295-301. [doi: 10.1016/j.pec.2008.11.015] [Medline: 19150199]
13. Sabaté E, World Health Organization. Adherence to Long-term Therapies: Evidence for Action. Geneva, Switzerland: World Health Organization; 2003.
14. Castro EM, Van Regenmortel T, Vanhaecht K, Sermeus W, Van Hecke A. Patient empowerment, patient participation and patient-centeredness in hospital care: a concept analysis based on a literature review. Patient Educ Couns 2016 Dec;99(12):1923-1939. [doi: 10.1016/j.pec.2016.07.026] [Medline: 27450481]
15. Lutfey KE, Wishner WJ. Beyond "compliance" is "adherence". Improving the prospect of diabetes care. Diabetes Care 1999 Apr;22(4):635-639. [doi: 10.2337/diacare.22.4.635] [Medline: 10189544]
16. Krass I, Schieback P, Dhippayom T. Adherence to diabetes medication: a systematic review. Diabet Med 2015 Jun;32(6):725-737. [doi: 10.1111/dme.12651] [Medline: 25440507]
17. Nieuwlaat R, Wilczynski N, Navarro T, Hobson N, Jeffery R, Keepanasseril A, et al. Interventions for enhancing medication adherence. Cochrane Database Syst Rev 2014 Nov 20;2014(11):CD000011 [FREE Full text] [doi: 10.1002/14651858.CD000011.pub4] [Medline: 25412402]
18. Rosli MR, Neoh CF, Wu DB, Hassan NW, Mahmud M, Rahimi A, et al. Evaluation of home medication review for patients with type 2 diabetes mellitus by community pharmacists: a randomised controlled trial. Pharm Pract (Granada) 2021;19(3):2397 [FREE Full text] [doi: 10.18549/PharmPract.2021.3.2397] [Medline: 34621450]
19. Kandula NR, Malli T, Zei CP, Larsen E, Baker DW. Literacy and retention of information after a multimedia diabetes education program and teach-back. J Health Commun 2011;16 Suppl 3:89-102. [doi: 10.1080/10810730.2011.604382] [Medline: 21951245]
20. Kravitz RL, Hays RD, Sherbourne CD. Recall of recommendations and adherence to advice among patients with chronic medical conditions. Arch Intern Med 1993 Aug 23;153(16):1869-1878. [doi: 10.1001/archinte.1993.00410160029002]
21. van Weert JC, Jansen J, Spreeuwenberg PM, van Dulmen S, Bensing JM. Effects of communication skills training and a question prompt sheet to improve communication with older cancer patients: a randomized controlled trial. Crit Rev Oncol Hematol 2011 Oct;80(1):145-159. [doi: 10.1016/j.critrevonc.2010.10.010] [Medline: 21075644]
22. Fallows RR, Hilsabeck RC. Comparing two methods of delivering neuropsychological feedback. Arch Clin Neuropsychol 2013 Mar;28(2):180-188. [doi: 10.1093/arclin/acs142] [Medline: 23315402]
23. Meth M, Calamia M, Tranel D. Does a simple intervention enhance memory and adherence for neuropsychological recommendations? Appl Neuropsychol Adult 2016;23(1):21-28. [doi: 10.1080/23279095.2014.996881] [Medline: 26111163]
24. Sherlock A, Brownie S. Patients' recollection and understanding of informed consent: a literature review. ANZ J Surg 2014 Apr;84(4):207-210. [doi: 10.1111/ans.12555] [Medline: 24812707]
25. Castro CM, Wilson C, Wang F, Schillinger D. Babel babble: physicians' use of unclarified medical jargon with patients. Am J Health Behav 2007;31 Suppl 1:S85-S95. [doi: 10.5555/ajhb.2007.31.supp.S85] [Medline: 17931142]

XSL•FO
RenderX

26. Schillinger D, Piette J, Grumbach K, Wang F, Wilson C, Daher C, et al. Closing the loop: physician communication with diabetic patients who have low health literacy. Arch Intern Med 2003 Jan 13;163(1):83-90. [doi: 10.1001/archinte.163.1.83] [Medline: 12523921]
27. McCarthy DM, Waite KR, Curtis LM, Engel KG, Baker DW, Wolf MS. What did the doctor say? Health literacy and recall of medical instructions. Med Care 2012 Apr;50(4):277-282 [FREE Full text] [doi: 10.1097/MLR.0b013e318241e8e1] [Medline: 22411440]
28. Tan SS, Goonawardene N. Internet health information seeking and the patient-physician relationship: a systematic review. J Med Internet Res 2017 Jan 19;19(1):e9 [FREE Full text] [doi: 10.2196/jmir.5729] [Medline: 28104579]
29. Ciciriello S, Johnston RV, Osborne RH, Wicks I, deKroo T, Clerehan R, et al. Multimedia educational interventions for consumers about prescribed and over-the-counter medications. Cochrane Database Syst Rev 2013 Apr 30(4):CD008416. [doi: 10.1002/14651858.CD008416.pub2] [Medline: 23633355]
30. Sim V, Galbraith K. Effectiveness of multimedia interventions in the provision of patient education on anticoagulation therapy: a review. Patient Educ Couns 2020 Oct;103(10):2009-2017. [doi: 10.1016/j.pec.2020.05.003] [Medline: 32532633]
31. Ab Malik N, Zhang J, Lam OL, Jin L, McGrath C. Effectiveness of computer-aided learning in oral health among patients and caregivers: a systematic review. J Am Med Inform Assoc 2017 Jan;24(1):209-217 [FREE Full text] [doi: 10.1093/jamia/ocw045] [Medline: 27274013]
32. Kuwabara A, Su S, Krauss J. Utilizing digital health technologies for patient education in lifestyle medicine. Am J Lifestyle Med 2019 Dec 13;14(2):137-142 [FREE Full text] [doi: 10.1177/1559827619892547] [Medline: 32231478]
33. Nealon J, Moreno A. Agent-based applications in health care. In: Moreno A, Nealon JL, editors. Applications of Software Agent Technology in the Health Care Domain. Basel, Switzerland: Birkhäuser; 2003.
34. Niederman F, Baker EW. Ethics and AI issues: old container with new wine? Inf Syst Front 2022 Jun 30;25:9-28. [doi: 10.1007/s10796-022-10305-1]
35. Feuerriegel S, Dolata M, Schwabe G. Fair AI. Bus Inf Syst Eng 2020 May 11;62:379-384. [doi: 10.1007/s12599-020-00650-3]
36. McMahan RP, Lai C, Pal SK. Interaction fidelity: the uncanny valley of virtual reality interactions. In: Proceedings of the 8th International Conference, VAMR 2016, Held as Part of HCI International 2016. 2016 Presented at: 8th International Conference, VAMR 2016, Held as Part of HCI International 2016; July 17-22, 2016; Toronto, ON. [doi: 10.1007/978-3-319-39907-2_6]
37. Mori M, MacDorman KF, Kageki N. The uncanny valley [from the field]. IEEE Robot Autom Mag 2012 Jun;19(2):98-100. [doi: 10.1109/MRA.2012.2192811]
38. Dolata M, Kilic M, Schwabe G. When a computer speaks institutional talk: exploring challenges and potentials of virtual assistants in face-to-face advisory services. In: Proceedings of the 52nd Hawaii International Conference on System Sciences (HICSS 2019). 2019 Presented at: HICSS 2019; January 8, 2019; Maui, Hawaii URL: https://www.zora.uzh.ch/id/eprint/173360/1/hicss_agent_v14_camera_zora.pdf [doi: 10.24251/hicss.2019.014]
39. Mazzi C, Ganguly P, Kidd M. Healthcare applications based on software agents. In: Proceedings of the 10th World Congress on Medical Informatics. 2001 Presented at: WCIM 2001; September 2-5, 2001; London, UK.
40. Linders GM, Vaitonytė J, Alimardani M, Mitev KO, Louwerse MM. A realistic, multimodal virtual agent for the healthcare domain. In: Proceedings of the 22nd ACM International Conference on Intelligent Virtual Agents. 2022 Presented at: IVA '22; September 6-9, 2022; Faro, Portugal. [doi: 10.1145/3514197.3551250]
41. Troy E, Doltani D, Harmon D. The role of a companion attending consultations with the patient. A systematic review. Ir J Med Sci 2019 Aug;188(3):743-750. [doi: 10.1007/s11845-018-1920-0] [Medline: 30374801]
42. Laidsaar-Powell RC, Butow PN, Bu S, Charles C, Gafni A, Lam WW, et al. Physician-patient-companion communication and decision-making: a systematic review of triadic medical consultations. Patient Educ Couns 2013 Apr;91(1):3-13. [doi: 10.1016/j.pec.2012.11.007] [Medline: 23332193]
43. Wolff JL, Roter DL. Family presence in routine medical visits: a meta-analytical review. Soc Sci Med 2011 Mar;72(6):823-831 [FREE Full text] [doi: 10.1016/j.socscimed.2011.01.015] [Medline: 21353358]
44. Bracher M, Stewart S, Reidy C, Allen C, Townsend K, Brindle L. Partner involvement in treatment-related decision making in triadic clinical consultations - a systematic review of qualitative and quantitative studies. Patient Educ Couns 2020 Feb;103(2):245-253. [doi: 10.1016/j.pec.2019.08.031] [Medline: 31477515]
45. Pearce C, Dwan K, Arnold M, Phillips C, Trumble S. Doctor, patient and computer--a framework for the new consultation. Int J Med Inform 2009 Jan;78(1):32-38. [doi: 10.1016/j.ijmedinf.2008.07.002] [Medline: 18752989]
46. Pearce C, Arnold M, Phillips C, Trumble S, Dwan K. The patient and the computer in the primary care consultation. J Am Med Inform Assoc 2011;18(2):138-142 [FREE Full text] [doi: 10.1136/jamia.2010.006486] [Medline: 21262923]
47. Mørck P, Langhoff TO, Christophersen M, Møller AK, Bjørn P. Variations in oncology consultations: how dictation allows variations to be documented in standardized ways. Comput Supported Coop Work 2018 May 18;27:539-568. [doi: 10.1007/s10606-018-9332-2]
48. Barrows RC, Clayton PD. Privacy, confidentiality, and electronic medical records. J Am Med Inform Assoc 1996;3(2):139-148 [FREE Full text] [doi: 10.1136/jamia.1996.96236282] [Medline: 8653450]
49. Keshta I, Odeh A. Security and privacy of electronic health records: concerns and challenges. Egypt Inform J 2021 Jul;22(2):177-183. [doi: 10.1016/j.eij.2020.07.003]

XSL•FO
RenderX

50. Pearce C, Sandoval M. Consulting with a computer: new frontiers. Aust J Gen Pract 2020 Sep;49(9):612-614 [FREE Full text] [doi: 10.31128/AJGP-05-20-5400] [Medline: 32864673]
51. Rose D, Richter LT, Kapustin J. Patient experiences with electronic medical records: lessons learned. J Am Assoc Nurse Pract 2014 Dec;26(12):674-680 [FREE Full text] [doi: 10.1002/2327-6924.12170] [Medline: 25234112]
52. Asan O, Young HN, Chewning B, Montague E. How physician electronic health record screen sharing affects patient and doctor non-verbal communication in primary care. Patient Educ Couns 2015 Mar;98(3):310-316 [FREE Full text] [doi: 10.1016/j.pec.2014.11.024] [Medline: 25534022]
53. Alkureishi MA, Lee WW, Lyons M, Press VG, Imam S, Nkansah-Amankra A, et al. Impact of electronic medical record use on the patient-doctor relationship and communication: a systematic review. J Gen Intern Med 2016 May;31(5):548-560 [FREE Full text] [doi: 10.1007/s11606-015-3582-1] [Medline: 26786877]
54. Gawande A. Why doctors hate their computers. The New Yorker. 2018 Nov 5. URL: https://www.newyorker.com/magazine/2018/11/12/why-doctors-hate-their-computers [accessed 2019-02-18]
55. Frankel R, Altschuler A, George S, Kinsman J, Jimison H, Robertson NR, et al. Effects of exam-room computing on clinician-patient communication: a longitudinal qualitative study. J Gen Intern Med 2005 Aug;20(8):677-682 [FREE Full text] [doi: 10.1111/j.1525-1497.2005.0163.x] [Medline: 16050873]
56. Kazmi Z. Effects of exam room EHR use on doctor-patient communication: a systematic literature review. Inform Prim Care 2013;21(1):30-39 [FREE Full text] [doi: 10.14236/jhi.v21i1.37] [Medline: 24629654]
57. Scalia P, Durand MA, Berkowitz JL, Ramesh NP, Faber MJ, Kremer JA, et al. The impact and utility of encounter patient decision aids: systematic review, meta-analysis and narrative synthesis. Patient Educ Couns 2019 May;102(5):817-841. [doi: 10.1016/j.pec.2018.12.020] [Medline: 30612829]
58. Stacey D, Légaré F, Lewis K, Barry MJ, Bennett CL, Eden KB, et al. Decision aids for people facing health treatment or screening decisions. Cochrane Database Syst Rev 2017 Apr 12;4(4):CD001431 [FREE Full text] [doi: 10.1002/14651858.CD001431.pub5] [Medline: 28402085]
59. Scalia P, Tighe KM, Elwyn G, Bagley PJ, Blunt HB, Boh B, et al. A systematic review of decision aids for gender affirming therapy. Transl Androl Urol 2021 Jun;10(6):2574-2582 [FREE Full text] [doi: 10.21037/tau-20-1000] [Medline: 34295744]
60. Agoritsas T, Heen AF, Brandt L, Alonso-Coello P, Kristiansen A, Akl EA, et al. Decision aids that really promote shared decision making: the pace quickens. BMJ 2015 Feb 10;350:g7624 [FREE Full text] [doi: 10.1136/bmj.g7624] [Medline: 25670178]
61. Elwyn G, Lloyd A, Joseph-Williams N, Cording E, Thomson R, Durand MA, et al. Option Grids: shared decision making made easier. Patient Educ Couns 2013 Feb;90(2):207-212. [doi: 10.1016/j.pec.2012.06.036] [Medline: 22854227]
62. Irizarry T, DeVito Dabbs A, Curran CR. Patient portals and patient engagement: a state of the science review. J Med Internet Res 2015 Jun 23;17(6):e148 [FREE Full text] [doi: 10.2196/jmir.4255] [Medline: 26104044]
63. Dameff C, Clay B, Longhurst CA. Personal health records: more promising in the smartphone era? JAMA 2019 Jan 29;321(4):339-340. [doi: 10.1001/jama.2018.20434] [Medline: 30633300]
64. Ammenwerth E, Schnell-Inderst P, Hoerbst A. The impact of electronic patient portals on patient care: a systematic review of controlled trials. J Med Internet Res 2012 Nov 26;14(6):e162 [FREE Full text] [doi: 10.2196/jmir.2238] [Medline: 23183044]
65. Goldzweig CL, Orshansky G, Paige NM, Towfigh AA, Haggstrom DA, Miake-Lye I, et al. Electronic patient portals: evidence on health outcomes, satisfaction, efficiency, and attitudes: a systematic review. Ann Intern Med 2013 Nov 19;159(10):677-687. [doi: 10.7326/0003-4819-159-10-201311190-00006] [Medline: 24247673]
66. mHealth economics 2017 – current status and future trends in mobile health. Research2Guidance. URL: https://research2guidance.com/product/mhealth-economics-2017-current-status-and-future-trends-in-mobile-health/ [accessed 2021-04-12]
67. Garabedian LF, Ross-Degnan D, Wharam JF. Mobile phone and smartphone technologies for diabetes care and self-management. Curr Diab Rep 2015 Dec;15(12):109 [FREE Full text] [doi: 10.1007/s11892-015-0680-8] [Medline: 26458380]
68. Holmen H, Wahl AK, Cvancarova Småstuen M, Ribu L. Tailored communication within mobile apps for diabetes self-management: a systematic review. J Med Internet Res 2017 Jun 23;19(6):e227 [FREE Full text] [doi: 10.2196/jmir.7045] [Medline: 28645890]
69. McMillan KA, Kirk A, Hewitt A, MacRury S. A systematic and integrated review of mobile-based technology to promote active lifestyles in people with type 2 diabetes. J Diabetes Sci Technol 2017 Mar;11(2):299-307 [FREE Full text] [doi: 10.1177/1932296816656018] [Medline: 27334301]
70. Nahum-Shani I, Smith SN, Spring BJ, Collins LM, Witkiewitz K, Tewari A, et al. Just-in-time adaptive interventions (JITAIs) in mobile health: key components and design principles for ongoing health behavior support. Ann Behav Med 2018 May 18;52(6):446-462 [FREE Full text] [doi: 10.1007/s12160-016-9830-8] [Medline: 27663578]
71. Eysenbach G. The law of attrition. J Med Internet Res 2005 Mar 31;7(1):e11 [FREE Full text] [doi: 10.2196/jmir.7.1.e11] [Medline: 15829473]

XSL•FO
RenderX

72. Fukuoka Y, Gay C, Haskell W, Arai S, Vittinghoff E. Identifying factors associated with dropout during prerandomization run-in period from an mHealth physical activity education study: the mPED trial. JMIR Mhealth Uhealth 2015 Apr 13;3(2):e34 [FREE Full text] [doi: 10.2196/mhealth.3928] [Medline: 25872754]
73. Lobelo F, Kelli HM, Tejedor SC, Pratt M, McConnell MV, Martin SS, et al. The wild wild west: a framework to integrate mHealth software applications and wearables to support physical activity assessment, counseling and interventions for cardiovascular disease risk reduction. Prog Cardiovasc Dis 2016;58(6):584-594 [FREE Full text] [doi: 10.1016/j.pcad.2016.02.007] [Medline: 26923067]
74. Zhang Y, Li X, Luo S, Liu C, Liu F, Zhou Z. Exploration of users' perspectives and needs and design of a type 1 diabetes management mobile app: mixed-methods study. JMIR Mhealth Uhealth 2018 Sep 21;6(9):e11400 [FREE Full text] [doi: 10.2196/11400] [Medline: 30249580]
75. Gordon WJ, Landman A, Zhang H, Bates DW. Beyond validation: getting health apps into clinical practice. NPJ Digit Med 2020;3:14 [FREE Full text] [doi: 10.1038/s41746-019-0212-z] [Medline: 32047860]
76. Heidel A, Hagist C. Potential benefits and risks resulting from the introduction of health apps and wearables into the German statutory health care system: scoping review. JMIR Mhealth Uhealth 2020 Sep 23;8(9):e16444 [FREE Full text] [doi: 10.2196/16444] [Medline: 32965231]
77. Street RLJ, Liu L, Farber NJ, Chen Y, Calvitti A, Zuest D, et al. Provider interaction with the electronic health record: the effects on patient-centered communication in medical encounters. Patient Educ Couns 2014 Sep;96(3):315-319 [FREE Full text] [doi: 10.1016/j.pec.2014.05.004] [Medline: 24882086]
78. Cahn A, Akirov A, Raz I. Digital health technology and diabetes management. J Diabetes 2018 Jan;10(1):10-17. [doi: 10.1111/1753-0407.12606] [Medline: 28872765]
79. Nurgalieva L, O'Callaghan D, Doherty G. Security and privacy of mHealth applications: a scoping review. IEEE Access 2020;8:104247-104268. [doi: 10.1109/access.2020.2999934]
80. Baker DW, Wolf MS, Feinglass J, Thompson JA, Gazmararian JA, Huang J. Health literacy and mortality among elderly persons. Arch Intern Med 2007 Jul 23;167(14):1503-1509. [doi: 10.1001/archinte.167.14.1503] [Medline: 17646604]
81. Magnani JW, Mujahid MS, Aronow HD, Cené CW, Dickson VV, Havranek E, et al. Health literacy and cardiovascular disease: fundamental relevance to primary and secondary prevention: a scientific statement from the American Heart Association. Circulation 2018 Jul 10;138(2):e48-e74 [FREE Full text] [doi: 10.1161/CIR.0000000000000579] [Medline: 29866648]
82. Michou M, Panagiotakos DB, Costarelli V. Low health literacy and excess body weight: a systematic review. Cent Eur J Public Health 2018 Sep;26(3):234-241 [FREE Full text] [doi: 10.21101/cejph.a5172] [Medline: 30419628]
83. N'Goran AA, Pasquier J, Deruaz-Luyet A, Burnand B, Haller DM, Neuner-Jehle S, et al. Factors associated with health literacy in multimorbid patients in primary care: a cross-sectional study in Switzerland. BMJ Open 2018 Feb 13;8(2):e018281 [FREE Full text] [doi: 10.1136/bmjopen-2017-018281] [Medline: 29440210]
84. Cramer JA. Consequences of intermittent treatment for hypertension: the case for medication compliance and persistence. Am J Manag Care 1998 Nov;4(11):1563-1568 [FREE Full text] [Medline: 10338902]
85. Collins LM. Optimization of Behavioral, Biobehavioral, and Biomedical Interventions: The Multiphase Optimization Strategy (MOST). Cham, Switzerland: Springer International Publishing; 2018.
86. Hardeman W, Houghton J, Lane K, Jones A, Naughton F. A systematic review of just-in-time adaptive interventions (JITAIs) to promote physical activity. Int J Behav Nutr Phys Act 2019 Apr 03;16(1):31 [FREE Full text] [doi: 10.1186/s12966-019-0792-7] [Medline: 30943983]
87. Dahodwala M, Geransar R, Babion J, de Grood J, Sargious P. The impact of the use of video-based educational interventions on patient outcomes in hospital settings: a scoping review. Patient Educ Couns 2018 Dec;101(12):2116-2124. [doi: 10.1016/j.pec.2018.06.018] [Medline: 30087021]
88. Miller CK. Adaptive intervention designs to promote behavioral change in adults: what is the evidence? Curr Diab Rep 2019 Jan 25;19(2):7. [doi: 10.1007/s11892-019-1127-4] [Medline: 30684109]
89. Gross T, Koch M. Computer-supported cooperative work - concepts and trends. In: Proceedings of the 11th International Conference of the Association Information and Management (AIM). 2006 Presented at: AIM 2006; June 8-9, 2006; Luxembourg.
90. Wilson P. Computer supported cooperative work: an overview. Intell Tutor Media 1990 Jan;1(3):103-116. [doi: 10.1080/14626269009408268]
91. Grudin J. Computer-supported cooperative work: history and focus. Computer 1994 May;27(5):19-26. [doi: 10.1109/2.291294]
92. Carroll JM, Rosson MB, Convertino G, Ganoe CH. Awareness and teamwork in computer-supported collaborations. Interact Comput 2006 Jan;18(1):21-46. [doi: 10.1016/j.intcom.2005.05.005]
93. Dolata M, Schwabe G. Paper practices in institutional talk: how financial advisors impress their clients. Comput Support Coop Work 2017 Jun 16;26(4-6):769-805. [doi: 10.1007/s10606-017-9279-8]
94. Willis M. I'm trying to find my way of staying organized: the socio-technical assemblages of personal health information management. Comput Support Coop Work 2019 May 15;28(6):1073-1102. [doi: 10.1007/s10606-019-09346-z]
95. Stebbins RA. Exploratory Research in the Social Sciences. Thousand Oaks, CA: SAGE Publications; 2001.

XSL•FO
RenderX

96. Nicholls CM, Lewis J, Ritchie J, Ormston R. Qualitative Research Practice: A Guide for Social Science Students and Researchers. Thousand Oaks, CA: SAGE Publications; 2003.
97. Färber A, Zigan N, Dolata M, Stalder P, Koppitz A, Schwabe G. The digital transformation of physician–patient consultations: identifying problems and approaches to improve adherence. In: Proceedings of the 52nd Hawaii International Conference on System Sciences. 2019 Presented at: HICSS 2019; January 8-11, 2019; Maui, Hawaii. [doi: 10.24251/hicss.2019.500]
98. Merriam SB, Grenier RS. Qualitative Research in Practice: Examples for Discussion and Analysis. Hoboken, NJ: John Wiley & Sons; 2019.
99. Tong A, Sainsbury P, Craig J. Consolidated criteria for reporting qualitative research (COREQ): a 32-item checklist for interviews and focus groups. Int J Qual Health Care 2007 Dec;19(6):349-357. [doi: 10.1093/intqhc/mzm042] [Medline: 17872937]
100. World Medical Association. World Medical Association Declaration of Helsinki: ethical principles for medical research involving human subjects. JAMA 2013 Nov 27;310(20):2191-2194. [doi: 10.1001/jama.2013.281053] [Medline: 24141714]
101. Palinkas LA, Horwitz SM, Green CA, Wisdom JP, Duan N, Hoagwood K. Purposeful sampling for qualitative data collection and analysis in mixed method implementation research. Adm Policy Ment Health 2015 Sep;42(5):533-544 [FREE Full text] [doi: 10.1007/s10488-013-0528-y] [Medline: 24193818]
102. Saunders B, Sim J, Kingstone T, Baker S, Waterfield J, Bartlam B, et al. Saturation in qualitative research: exploring its conceptualization and operationalization. Qual Quant 2018;52(4):1893-1907 [FREE Full text] [doi: 10.1007/s11135-017-0574-8] [Medline: 29937585]
103. Gordon AV, Rohrbeck R, Schwarz JO. Escaping the ‘faster horses’ trap: bridging strategic foresight and design-based innovation. Technol Innov Manag Rev 2019 Sep;9(8):30-42 [FREE Full text] [doi: 10.22215/timreview/1259]
104. Dolata M, Schwabe G. Design thinking in IS Research Projects. In: Brenner W, Uebernickel F, editors. Design Thinking for Innovation. Cham, Switzerland: Springer; 2016.
105. Mayring P. Qualitative Inhaltsanalyse. Grundlagen und Techniken. Weinheim, Germany: Deutscher Studien Verlag; 2000.
106. Birjali M, Kasri M, Beni-Hssane A. A comprehensive survey on sentiment analysis: approaches, challenges and trends. Knowl Based Syst 2021 Aug;226:107134. [doi: 10.1016/j.knosys.2021.107134]
107. Schnitman G, Wang T, Kundu S, Turkdogan S, Gotlieb R, How J, et al. The role of digital patient education in maternal health: a systematic review. Patient Educ Couns 2022 Mar;105(3):586-593. [doi: 10.1016/j.pec.2021.06.019] [Medline: 34183217]
108. Färber A, de Spindler A, Moser A, Schwabe G. Closing the loop for patients with chronic diseases - from problems to a solution architecture. In: Proceedings of the IEEE 11th International Conference on Healthcare Informatics (ICHI). 2023 Presented at: ICHI 2023; June 26-29, 2023; Houston, TX, USA URL: https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=10337012 [doi: 10.1109/ICHI57859.2023.00055]
109. Hämäläinen A, Hirvonen H. Electronic health records reshaping the socio-technical practices in long-term care of older persons. Technol Soc 2020 Aug;62:101316. [doi: 10.1016/j.techsoc.2020.101316]
110. Janett RS, Yeracaris PP. Electronic medical records in the American health system: challenges and lessons learned. Cien Saude Colet 2020 Mar;25(4):1293-1304 [FREE Full text] [doi: 10.1590/1413-81232020254.28922019] [Medline: 32267432]
111. Wisner K, Lyndon A, Chesla CA. The electronic health record's impact on nurses' cognitive work: an integrative review. Int J Nurs Stud 2019 Jun;94:74-84. [doi: 10.1016/j.ijnurstu.2019.03.003] [Medline: 30939418]

Abbreviations

AI: artificial intelligence
COREQ: Consolidated Criteria for Reporting Qualitative Research
CSCW: computer-supported cooperative work
DC: digital companion
EMR: electronic medical record
mHealth: mobile health
PDA: patient decision aid
PHR: personal health record

XSL•FO
RenderX

Edited by B Price; submitted 04.06.23; peer-reviewed by L Weinert, Y Zhang, C Yu, C Baur; comments to author 04.08.23; revised version received 09.12.23; accepted 15.01.24; published 18.03.24.

Please cite as:
Färber A, Schwabe C, Stalder PH, Dolata M, Schwabe G
Physicians' and Patients' Expectations From Digital Agents for Consultations: Interview Study Among Physicians and Patients
JMIR Hum Factors 2024;11:e49647
URL: https://humanfactors.jmir.org/2024/1/e49647
doi:10.2196/49647
PMID:38498022

©Andri Färber, Christiane Schwabe, Philipp H Stalder, Mateusz Dolata, Gerhard Schwabe. Originally published in JMIR Human Factors (https://humanfactors.jmir.org), 18.03.2024. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on https://humanfactors.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Preferences for mHealth Intervention to Address Mental Health Challenges Among Men Who Have Sex With Men in Nepal: Qualitative Study

Kamal Gautam[1*], MPH; Camille Aguilar[2*]; Kiran Paudel[1,3], BPH; Manisha Dhakal[4], MBS; Jeffrey A Wickersham[5], PhD; Bibhav Acharya[6,7], MD, PhD; Sabitri Sapkota[7], MPH, PhD; Keshab Deuba[8], PhD; Roman Shrestha[1,5], MPH, PhD

[1]Department of Allied Health Sciences, University of Connecticut, Storrs, CT, United States
[2]School of Nursing and Health Studies, University of Miami, Miami, FL, United States
[3]Nepal Health Frontiers, Tokha-5, Kathmandu, Nepal
[4]Blue Diamond Society, Dhumbarahi Height, Kathmandu, Nepal
[5]Department of Internal Medicine, Section of Infectious Diseases, Yale School of Medicine, New Haven, CT, United States
[6]Department of Psychiatry and Behavioral Sciences, School of Medicine, UCSF Weill Institute for Neurosciences, San Francisco, CA, United States
[7]Possible, Kathmandu, Nepal
[8]Department of Global Public Health, Karolinska Institutet, Stockholm, Sweden
[*]these authors contributed equally

Corresponding Author:
Roman Shrestha, MPH, PhD
Department of Allied Health Sciences
University of Connecticut
358 Mansfield Road
Storrs, CT, 06269
United States
Phone: 1 860 486 2834
Email: roman.shrestha@uconn.edu

Abstract

Background: Men who have sex with men (MSM) are disproportionately burdened by poor mental health. Despite the increasing burden, evidence-based interventions for MSM are largely nonexistent in Nepal.

Objective: This study explored mental health concerns, contributing factors, barriers to mental health care and support, and preferred interventions to improve access to and use of mental health support services among MSM in Nepal.

Methods: We conducted focus groups with MSM in Kathmandu, Nepal, in January 2023. In total, 28 participants took part in 5 focus group sessions. Participants discussed several topics related to the mental health issues they experienced, factors contributing to these issues, and their suggestions for potential interventions to address existing barriers. The discussions were recorded, transcribed, and analyzed using Dedoose (version 9.0.54; SocioCultural Research Consultants, LLC) software for thematic analysis.

Results: Participants reported substantial mental health problems, including anxiety, depression, suicidal ideation, and behaviors. Contributing factors included family rejection, isolation, bullying, stigma, discrimination, and fear of HIV and other sexually transmitted infections. Barriers to accessing services included cost, lack of lesbian, gay, bisexual, transgender, intersex, queer, and asexual (LGBTIQA+)–friendly providers, and the stigma associated with mental health and sexuality. Participants suggested a smartphone app with features such as a mental health screening tool, digital consultation, helpline number, directory of LGBTIQA+-friendly providers, mental health resources, and a discussion forum for peer support as potential solutions. Participants emphasized the importance of privacy and confidentiality to ensure mobile apps are safe and accessible.

Conclusions: The findings of this study have potential transferability to other low-resource settings facing similar challenges. Intervention developers can use these findings to design tailored mobile apps to facilitate mental health care delivery and support for MSM and other marginalized groups.

XSL•FO
RenderX

(JMIR Hum Factors 2024;11:e56002) doi:10.2196/56002

KEYWORDS

mental health; MSM; mHealth; smartphone apps; digital health; Nepal; gay; homosexual; homosexuality; men who have sex with men; focus group; focus groups; qualitative; barrier; barriers; thematic; mHealth; mobile health; app; apps; applications; applications

Introduction

Gay, bisexual, and other men who have sex with men (MSM) have poorer mental health and experience more mental distress than their cisgender heterosexual counterparts [1-3]. Studies have shown a high proportion of MSM's experiences such as mood swings, disordered eating behavior, anxiety disorder, depression, suicidal ideation and behaviors, substance abuse, and body image disorders [4-7]. A recent systematic review and meta-analysis found that the prevalence of depression among MSM in Asia was 37% [6]. These mental health issues experienced by MSM are often linked to stressors triggered by a homophobic environment, particularly due to their sexual orientation [8].

In the context of Nepal, homosexuality is not criminalized, and the rights of MSM are guaranteed by the constitution [9,10]. Despite these legal safeguards, the prevailing cultural norms and societal attitudes pose significant challenges. Traditional and cultural values emphasize heterosexual marriages and family structures and traditional expectations of relationships, and a lack of family support often marginalizes individuals with diverse sexual orientations [11]. These social and cultural characteristics create a heteronormative and stigmatizing environment for MSM, which is detrimental to their mental health. Past studies have found that a very high number of MSM in Nepal had clinically significant depression (54%) and lifetime prevalence of suicidal thoughts (26%) [12,13]. Despite these dire mental health statistics, MSM encounter barriers in accessing health care, particularly mental health services, due to social stigma, discrimination, financial constraints, and insensitivity among health care providers [11,12,14-17]. These barriers to seeking mental health and psychosocial support among MSM, who not only have the highest needs but also the highest unmet needs, give rise to health disparities in this population. In order to reduce these disparities, improving access is crucial for advancing their overall health and well-being.

Mobile health (mHealth), especially mobile apps, offers a promising solution to bridge this gap. It can offer tailored and cost-effective interventions without the need for in-person contact and can provide convenience, improve mental health literacy and easy accessibility, eliminate travel hassles, and encourage help-seeking behavior [18,19]. With Nepal experiencing significant growth in mobile phone ownership of 96% and over 70% using the internet through smartphones, mobile app–based interventions tailored to the needs of MSM in Nepal are potentially feasible [20]. Recognizing the potential of mHealth, we conducted this study to (1) identify the mental health challenges and barriers to accessing mental health and psychosocial support services among MSM and (2) understand their preferences for smartphone apps (eg, functionality, format, design, and attributes) that could enable their access to mental health and psychosocial support services access.

Methods

Study Setting and Recruitment

This qualitative study is part of a larger HIV biobehavioral survey that was conducted among 250 MSM participants in Kathmandu, Nepal [13]. Five focus group (FG) sessions were conducted with MSM participants in January 2023. Four of these sessions included 6 MSM participants in each, while the remaining session had 4 participants (N=28). FG sessions were conducted until a point of theoretical saturation was achieved. Eligibility criteria for participation included: (1) 18 years or older, (2) self-identified as cisgender MSM, and (3) proficiency in Nepali or English.

Participants were recruited using respondent-driven sampling, a network-based sampling method often used for hard-to-reach populations. The recruitment chain was initiated with 5 MSM "seeds," purposively selected based on recommendations from a community-based organization providing services to MSM. Each seed who completed the interviewer-administered questionnaire was given 5 recruitment coupons to recruit potential peers. Subsequent participants were, in turn, given 5 coupons to recruit additional peers. In total, 28 (~11%) of the survey participants were randomly selected for the FG sessions.

Study Procedure

FG sessions were conducted inside the community-based organization's office and lasted about 90 minutes. A semistructured FG topic guide with appropriate probes was developed that guided the discussion. A trained facilitator led the FG sessions, and a cofacilitator took the notes. Both the facilitator and cofacilitator identified themselves as MSM.

Before the discussion, participants completed an interviewer-administered Qualtrics survey that included sociodemographic, sexual health, alcohol, smoking, violence, and mental health–related questions. The participants' exposure to violence was assessed using the 4-item Hurt, Insult, Threat, and Scream screening tool, using a 5-point frequency format (scores 4-20). Final scores were classified as normal (0-10) or violence (11-25) [21]. Depressive symptoms were evaluated with the Patient Health Questionnaire instrument, scoring each of the 9 *Diagnostic and Statistical Manual of Mental Disorders*, 4th edition, criteria (0-3). A composite score of 0-27 was computed, with a score exceeding 10 indicating moderate to severe depressive symptoms [22].

The FGs involved questions and discussions about traumatic life events. Participants were made aware that they did not have to answer any questions that they felt were distressing and could leave the FG session at any time if they felt uncomfortable. A

XSL•FO
RenderX

study team member was also present at all 5 FG sessions to refer to a counselor or provide any additional support needed in the case of a distressing situation. While conducting the FG sessions, a trained facilitator approached participants sensitively, respecting moments of silence and their willingness to continue discussions—statements like “I am fine” or “we can continue” followed silence. Despite the sensitive topics discussed, none of the participants requested support, including speaking with counselors. At the end of all FG sessions, participants also disclosed that they were glad to have had the opportunity to share their experiences.

Data Analysis

SPSS (version 29.0.0 software; IBM Corp) was used to calculate descriptive statistics (frequencies and percentages) for the variables collected via a Qualtrics survey. FG transcripts were transcribed and checked for accuracy before coding. The 2 coders (KG and CA) read and reread transcriptions to identify key ideas and recurring themes. A codebook was developed with mutually agreed-upon codes derived from the FG transcripts, and coding was completed independently by 2 researchers (KG and CA). To ensure reliability, codes were constantly compared for agreement and discussed between the coders, and the senior author (RS) cross-checked all codes. Dedoose (version 9.0.54) was used for data management and analysis. The themes were gathered as child codes and then placed into a broad category as root codes. Each theme with its qualitative quotes to best illustrate the findings are presented in the results section.

Ethical Considerations

The study protocol was approved by the institutional review boards at the University of Connecticut (H22-0039) and the Nepal Health Research Council (2391-2022 P). All the participants provided verbal informed consent before their participation. Participants were explained the importance of maintaining the confidentiality of FGs and requested not to discuss the experiences and comments shared during the FGs with others. All the sessions were conducted in Nepali and were audio recorded, transcribed, and translated. Participants were compensated NRs 1000 (~US $8) for their time and participation. FG transcripts were deidentified before the analysis, and the survey data were anonymous.

Results

Participant Characteristics

Table 1 provides information on participants’ characteristics. The mean age of study participants was 25.3 (SD 6.1) years. Most of the 28 participants were Hindu (n=22, 79%), had a high school or higher degree (n=21, 75%), and identified as gay (n=22, 79%). A total of 21% (n=6) of participants had depressive symptoms, and 14% (n=4) had experienced violence in their life. A little over half (n=15, 54%) of participants had used health-related mobile apps, and almost 90% (n=25) used digital devices to search for health-related information.

XSL•FO
RenderX

Table 1. Participant characteristics (N=28).

Sociodemographic factors	Values
Age (years), mean (SD)	25.3 (6.1)
Religion, n (%)	
Hindu	22 (79)
Buddhist	5 (18)
Others	1 (4)
Level of education, n (%)	
Up to grade 10	7 (25)
High school and above	21 (75)
Employment, n (%)	
No	15 (54)
Yes	13 (46)
Income level, n (%)	
Less than NRs 20,000 (~US $150)	12 (43)
NRs 20,000 (~US $150) and above	16 (57)
Sexual orientation, n (%)	
Gay	22 (77)
Bisexual	6 (21)
Relationship status, n (%)	
Single	19 (68)
With partner	9 (32)
Depressive symptoms, n (%)	
No	22 (79)
Yes	6 (21)
Ever experienced violence, n (%)	
No	24 (86)
Yes	4 (14)
Daily smoker, n (%)	
No	5 (18)
Yes	23 (82)
Alcohol use (past 12 months), n (%)	
No	6 (21)
Yes	12 (79)
HIV status, n (%)	
Positive	1 (4)
Negative	27 (96)
Syphilis status, n (%)	
Positive	8 (29)
Negative	20 (71)
Engaged in anal sex (past 6 months), n (%)	
No	6 (21)
Yes	22 (79)
Condomless sex (past 6 months), n (%)	

XSL•FO
RenderX

Sociodemographic factors	Values
No	11 (39)
Yes	17 (61)
Sexual partners in (past 6 months), n (%)	
Single	17 (61)
Multiple	11 (39)
Engagement in group sex (past 6 months), n (%)	
No	26 (93)
Yes	2 (7)
Engagement in sex work (past 6 months), n (%)	
No	26 (93)
Yes	2 (7)
Has any health insurance, n (%)	
No	24 (86)
Yes	4 (14)
Use of health-related apps in mobile, n (%)	
No	13 (46)
Yes	15 (54)
Use of mobile or technological devices to search for health-related information, n (%)	
No	3 (11)
Yes	25 (89)

FG Results

Overview

Throughout the data analysis, 3 overarching themes emerged in the codebook with their own subthemes (Multimedia Appendix 1): (1) mental health challenges, (2) barriers to accessing mental health services, and (3) preference for mental health mobile apps with desired features and attributes.

Mental Health Challenges

Mental health challenges faced by the participants involve a multifaceted interaction of factors, including sexual orientation, emotional distress, stigma, discrimination and victimization, and social exclusion. Moreover, they frequently encounter barriers to accessing support services that could enhance their mental well-being. Participants not only vividly described their day-to-day challenges but also shared insights into the collective experiences of the lesbian, gay, bisexual, transgender, intersex, queer, and asexual (LGBTIQA+) community. Their comprehensive perspective underscored the profound impact of prevailing societal biases on their mental well-being (Textbox 1).

A constant fear of societal judgment and family pressure to conform to traditional gender norms has intensified issues like anxiety, depression, and suicidal thoughts. Participants in all FGs highlighted the pressure to enter heterosexual marriages, causing emotional turmoil as they navigate their identities and societal expectations (Textbox 2).

Participants disclosed coping mechanisms, such as drug use, drinking alcohol, smoking, engaging in sexual risk behaviors (eg, multiple sex partners), and self-harm. These strategies were described as providing temporary relief from the immense emotional turmoil they experience.

> *I have personally known someone who started risky sexual behavior from a young age because that was how they felt validated. They wanted others to make them feel better. So, they would often engage in multiple sexual encounters, thinking it would help them cope with their struggles... I also know people who turned to drugs and alcohol to cope with themselves.* [18-year-old participant from FG1]

> *…due to tension and mental pressure, it was tough for me to control myself, so I started to cut my hands with a razor; I did it many times. I was also thinking of taking tablets for suicide.* [30-year-old participant from FG3]

Several participants talked about and shared their experiences of intense anxiety and fear surrounding the possibility of contracting HIV and other sexually transmitted infections following sexual encounters with their partners.

> *I have extreme fear about whether I contracted it [HIV] or not… even the close friends I know have contracted HIV, and because of that, I also have a fear and anxiety of whether I contracted HIV or not after the sex is done.* [21-year-oldparticipant from FG1]

XSL•FO
RenderX

Textbox 1. Social acceptance and lack of family support heighten mental health challenges.

- "…there is no one, and when we open up there is no family support. Family supporting queer people, it is like gold which is rare. We only open out in this [LGBTIQA+] community; You can imagine how bad is our mental status and the situation." [26-year-old participant from FG2]
- "…because of my sexuality, sometimes I suffer from social anxiety, 'are they judging me because of my looks, voice or the way I dress." [24-years-old participant from FG5]
- "I study in 12th grade, and most of the time, I am bullied by my male classmates… even the teachers ask, "Why do you act like a girl?" And most of them do know I use TikTok, and everyone knows about me, so I think bullying is also another part, and I think mental health or stress is a common occurrence for everyone in LGBTIQA+ people." [22-year-old participant from FG1]

Textbox 2. Mental health challenges from societal pressure and identity concealment.

- "…if I have to be 'me' or do something feminine, then there is a fear of being judged by other people, so I have to pretend as a closeted man, I have to pretend masculine, have a masculine voice, and all the stereotypes and stigmas the people in the community have kept, which fuels the anxiety." [22-year-old participant from FG1]
- "I faced immense pressure from my family about marriage, being the only son. I kept my sexual orientation hidden, making it harder in our community. The talks of marriage became unbearable… I felt so distressed that I left home and was even suicidal. My relatives found out, but the misunderstanding about my identity remained... these struggles took a toll on my mental health, forcing me to search for ways to cope and maintain my well-being." [28-year-old participant from FG3]
- "I was so low that all I had in mind was suicide." [29-year-old participant from FG2]

Participants also shared that they fear the potential disclosure of their HIV status because they anticipate that others may treat them differently after learning their status.

> *When it comes to HIV, if a person status has HIV, he is afraid that if his status leaks, then people will look differently.* [23-year-old participant from FG4]

One participant recounted how their colleague, upon learning their HIV-positive status, tragically died by suicide, underscoring the emotional toll and mental health challenges.

> *There was one colleague of mine who died by suicide as soon as the HIV test result came back positive.* [25-year-old participant from FG1]

Barriers to Accessing Mental Health Services

Participants shared that many gay, bisexual, and MSM do not seek mental health services because they perceive themselves as mentally healthy and believe their lives are going well, leading them to overlook the need for such support.

> *The reason I believe that our community members do not seek mental health services is because they think they are alright, that their life is just going on, they think they are fine and healthy and feel they don't need such services.* [25-year-old participant from FG3]

Participants also shared that individuals tend to become more open and willing to seek help if they are aware of mental health services like counseling and therapy.

> *…if people are aware of counseling and therapy, people will be more willing to go there.* [24-year-old participant from FG2]

Participants discussed that individuals still closeted about their identity find it challenging to trust others, creating a communication barrier. Their hesitancy to trust stems from a history of hiding aspects of themselves, hindering open communication and sharing true feelings and experiences.

> *It is hard for people to trust. There is also a communication barrier because they are still closeted and grew up hiding things from the beginning. If the person themselves is not trusting them, then how can they trust the person in front of them.* [35-year-old participant from FG2]

Stigma and discrimination associated with mental health and sexuality were major concerns for participants. Many participants brought up fears of being labeled "*pagal*" (a pejorative that is closest to "crazy" in English) as a barrier to accessing mental health services.

> *…there is a stigma against mental health, that is the reason we do not seek mental health services. If we visit a health care center, then people will talk about it, and the peer groups and society will think of us as pagal (crazy); they will say that this person is taking medicines, so that is another reason we do not visit mental health care centers.* [26-year-old participant from FG3]

Others discussed the impact that homophobia can have on MSM seeking mental health services. Homophobia and heterosexism still exist in Nepal's society and can have significant impacts on MSM decisions.

> *A stereotypical saying "how can men like men?" is still prevalent in society, so, to not get judged by others, people don't attend these [mental health] sessions.* [29-year-old participant from FG2]

Many participants expressed their frustration with medical professionals who, instead of addressing their health concerns seriously, tend to label them, dismiss their issues, and attribute symptoms to perceived psychological factors such as overthinking, thereby hindering their access to necessary services.

> *Often, the doctor calls us with names, gives us a tag, they do not give us a priority, they only say "there's*

XSL•FO
RenderX

> *nothing wrong with you, it's only because you overthink" which has an impact on seeking services.* [26-year-old participant from FG3]

Many participants expressed the financial strain posed by mental health services for MSM in Nepal. The consensus was that the cost associated with psychiatrist visits, along with their limited financial resources, significantly affects MSMs' ability to access the necessary mental health services.

> *The main reason is finances because it is still very expensive, like we have to pay NRs 800 to 1000 (approximately US $8 to 10) per visit. It is expensive, even more so in private clinics.* [23-year-old participant from FG5]

Several participants across all FG sessions expressed concerns about time limitations and transportation challenges when it came to accessing mental health services. In an FG, a majority of participants agreed on the considerable difficulty that MSM faces in securing transportation to be able to go to a physical mental health appointment. In another session, everyone unanimously agreed and nodded in agreement with the following statement:

> *I think so too, because not everywhere has access to transportation, and for some places, we might even have to walk a lot to reach there.* [27-year-old participant from FG2]

When it comes to time constraints, participants talked about the difficulty of scheduling mental health appointments within the confines of work or school hours. They highlighted the difficulty of taking leave from work or school to attend counseling sessions during times of need.

> *I can go frequently, but the counseling appointment has to be time-friendly. Some of us are employed from 10 am to 5 pm or even 6 pm. If the counseling session is around that time, then I might come for a few days, taking a leave from work, but if my office does not allow me, then even if I had a mental health support need, I would not be able to attend.* [26-year-old participant from FG3]

Solution: Mental Health Smartphone Apps With Desired Features and Attributes

Overview

Participants expressed a preference for a smartphone app with a variety of features and attributes compared to traditional clinical settings. They foresaw that such an app could enhance understanding of mental health, offer convenience, improve accessibility, reduce the necessity for travel and associated expenses, and deliver services in a confidential and nonjudgmental setting.

> *During our young age, we didn't have any type of apps to help with our issues or any sort of networking apps like Grindr, but now people are more open to using apps, so creating an app to help solve the mental health issue and counsel can be a great idea.* [35-year-old participant from FG2]

Desired Features of the Mobile App

Participants recommended using creative approaches, such as fun activities to assess individuals' mental health for early detection, moving away from more direct approaches.

> *Something creative, not a direct approach, but through games or other ways we could assess the mental health status of the people for early detection.* [25-year-old participant from FG3]

Participants emphasized the importance of using the app to schedule regular counseling appointments with mental health professionals for those requiring assistance. There was a strong preference for using Zoom over platforms like Viber and WhatsApp for digital counseling, citing its widespread use during the COVID-19 pandemic.

> *…those who are in need of mental health services should get counseling appointments from a professional by selecting them once or twice a week in the app.* [29-year-old participants from FG4]

> *Rather than Viber and WhatsApp, Zoom is good for e-counseling, as in COVID many people are using it.* [18-year-old participants from FG4]

All participants underscored the importance of mental health and psychosocial service providers being qualified, friendly, and supportive of the LGBTIQA+ community. They stressed the need for an environment where individuals feel safe and comfortable to share their concerns.

> *First of all, they should be very friendly towards LGBTIQA+, no matter whether they are a community member or not, and we have to feel safe and able to share everything. Is qualified and has studied the related field.* [23-year-old participant from FG5]

Some participants had suggestions that would help make MSM more comfortable in participating in digital counseling, such as making cameras not compulsory.

> *We can do it through audio calls. Zoom counseling sessions are fine, but opening cameras should not be necessary or compulsory.* [21-year-old participant from FG5]

An additional recommendation included providing convenient hours, allowing users of the app to secure digital counseling appointments relatively quickly. This would accommodate individuals who work or go to school, ensuring continued accessibility to the services.

> *People will schedule according to their needs and how big their problem is, if you are having a problem now and get an appointment for a session after a month, it is not possible.* [21-year-old participant from FG2]

Participants suggested incorporating a toll-free helpline number within the smartphone app. They shared their experiences with toll-free helplines that did not function as intended in the past. Additionally, they provided suggestions for improving the toll-free helplines within the mobile app.

XSL•FO
RenderX

> *We can use a Toll-free helpline number, but even I tried to use toll-free service every time it was busy. So, the missed call system [call back system] is good.* [21-year-old participant from FG5]

Several participants suggested including a feature to message counselors in addition to the toll-free helpline that could help those who do not want to or cannot talk over the phone.

> *Some might not want to speak; they could talk through chat.* [30-year-old participant FG3]

Several participants shared their difficulties in finding friendly mental health and psychosocial service providers. To address this issue, participants suggested having a directory of LGBTIQA+-friendly providers on a mobile app that would help show MSM where to go when they require help.

> *I searched, and I came to know. It took me a lot of effort, and it was hard to find psychosocial counselors.* [29-year-old participant from FG2]

Participants also suggested to include mental health educational resources, especially in the form of videos.

> *Many are hidden, they do not even want to come out of the house, because of the fear of society. But they use mobile apps, they could connect to the app, and even with information and educational videos, we could reach them.* [21-year-old participant from FG5]

Many participants suggested a feature to connect with peers and other members of the MSM community through a communication channel within the app. They highlighted the importance of such a platform for sharing experiences and emphasized the value of peer support.

> *I think a discussion forum would be a good addition. The forum can help you share and make you feel like you are not the only one who is going through the same trauma and hardships, and we will be sharing with each other.* [21-year-old participant from FG1]

Attributes of the App

Many participants suggested placing special emphasis on the privacy and confidentiality of data collected by the app. They recommended that app developers and health care providers should commit to privacy and confidentiality clauses in their contracts, with strict consequences for any breaches of information.

> *The staff, app developers, and providers should sign on privacy and confidentiality in their contract. If leakage of information is found, they need to know that strong steps will be taken.* [26-year-old participant from FG3]

When discussing the user interface and colors of the app, several participants suggested that the mobile app should not overtly appear targeted exclusively at the LGBTIQA+ community. The participant expresses a desire for the app to have a discreet appearance, in contrast to the distinctiveness of dating apps targeted toward LGBTIQA+.

> *Through application maybe, the application should not look like for only LGBTIQ. It must look normal, not like Grindr.* [23-year-old participant from FG5]

Participants showed a strong interest in an engaging activity for user engagement and retention, particularly one that incorporates entertainment. One participant mentioned:

> *There has to be an environment in the mobile app so that I feel like going and using it again.* [26-year-old participant from FG3]

Discussion

Principal Findings

This study revealed a complex interplay between mental health challenges, including depression, anxiety, and suicidal behavior, among MSM in Nepal. The findings further highlight the barriers to accessing mental health care and support services among Nepali MSM due to factors such as insufficient mental health literacy, privacy concerns, financial strain, stigma, and discrimination. This underscores the urgent need for tailored and accessible mental health interventions. Participants overwhelmingly preferred smartphone app interventions to address the identified barriers and challenges, emphasizing their preference for accessible and confidential mental health support through digital platforms.

The major concern among MSM, where individuals perceive themselves as “all right” without the need for mental health services and less help-seeking attitudes, likely indicates a lack of mental health literacy, which is similar to the findings from studies among men and other minority populations [23-27]. Participants in this study expressed a preference for mental health resources and screening tools integrated into the app. Few studies have demonstrated that a smartphone app with an easily accessible and comprehensive mental health education module, resources, and engaging screening tools has the potential to combat this issue by fostering a proactive attitude toward mental well-being, the importance of seeking support, and the early detection of mental health problems [28-30].

The stigma and discrimination faced by MSM, both within society and health care settings, contribute to hesitancy in seeking mental health support. This fear of stigma and reluctance aligns with the findings from studies of various marginalized populations [26,31-33]. In response to this, participants expressed a preference for features within the mobile app that could link participants with LGBTIQA+-friendly mental health professionals through video sessions, automated text messages, or phone calls, emphasizing the crucial role of trust and understanding in the provider-patient relationship. Few interventions have integrated such features into digital interventions [34,35]. This feature could help to overcome this barrier by connecting individuals with LGBTIQA+-friendly and supportive mental health professionals and fostering a more inclusive, judgment-free, and accessible mental health support system.

In line with a substantial body of research, the findings emphasize that various stressors, particularly those related to societal biases, discrimination, fear of HIV, and other sexually

XSL•FO
RenderX

transmitted infection results, contribute to psychological distress, and these influence maladaptive coping behaviors among MSM [7,8,36-38]. By incorporating features such as mental health resources, coping strategies, and peer support discussion forums, the app can have the potential to empower MSM to navigate these challenges more effectively.

The privacy and confidentiality concerns expressed by MSM underscore the need for a sensitive approach to mental health support. This apprehension aligns with findings from studies of various minority populations [39-42]. Participants in this study articulated the desire for a mobile app that explicitly addresses these concerns through robust consent forms, privacy features, and secure messaging platforms. The app could have features that aim to ensure privacy and confidentiality, potentially fostering MSM trust and addressing barriers related to sharing personal mental health information. Integrating these features into app design could significantly contribute to alleviating privacy concerns and establishing a secure environment that encourages seeking mental health support.

The cost of accessing mental health services was a major concern for participants in the study, which aligns with previous research on the cost of mental health in Nepal [43]. It is important to address financial strain in any intervention that is created to help MSM in Nepal with mental health [44-46]. Studies have found that, by reducing travel expenses, mHealth interventions help allow access for sexual minority individuals to mental health care [47,48]. This not only addresses the financial challenges faced by Nepali MSM but also alleviates the transportation struggles [35].

Strengths and Limitations

This study is of particular value due to the lack of participant involvement in the development of mental health interventions, with LGBTIQA+ consultation being notably rare when it comes to the creation of health interventions, policies, or guidelines [49,50]. Using FGs, the participants' perspectives can be used to create a more tailored and effective digital health intervention. However, this study has its own limitations. One of these limitations is the presence of social desirability bias, which is a common occurrence in FG discussions. This bias can influence participants to express socially acceptable opinions rather than their true thoughts and feelings. Additionally, it is worth noting that the study was done in Kathmandu, Nepal, which can differ in culture and access to mental health services than other areas of Nepal, limiting the transferability of the study findings mainly on the challenges and barriers. Finally, it is important to consider that the desire to participate in a given intervention does not automatically guarantee its real-world adoption. Evaluating the actual usage and effectiveness of the intervention in real-life scenarios is crucial to fully understand its impact and potential benefits. Therefore, it is necessary to evaluate real-world usage.

Conclusions

The study highlights the mental health challenges encountered by MSM in Nepal and the barriers they face in accessing mental health support services. The participants' direct quote, "invisible in the corner of the room," captures the hidden nature of their struggles intimately tied to the intersectional stigma surrounding mental health and sexuality. Emphasizing the potential of mobile apps, our findings suggest that incorporating user-friendly features like accessible resources, mental health screening tools, and digital counseling with LGBTIQA+-friendly providers can bring visibility to the mental health challenges of MSM. The mobile app has the ability to establish an open and supportive space, breaking down barriers and offering a pathway for MSM in Nepal to identify and address their mental health concerns with ease and confidence.

Acknowledgments

RS received funding from the National Institute on Drug Abuse (award number K01DA051346).

Conflicts of Interest

RS is an editorial board member of JMIR mHealth and uHealth.

Multimedia Appendix 1
Parent and child codes with description.
[DOCX File , 17 KB - humanfactors_v11i1e56002_app1.docx]

References

1. Sivasubramanian M, Mimiaga MJ, Mayer KH, Anand VR, Johnson CV, Prabhugate P, et al. Suicidality, clinical depression, and anxiety disorders are highly prevalent in men who have sex with men in Mumbai, India: findings from a community-recruited sample. Psychol Health Med 2011 Aug;16(4):450-462 [FREE Full text] [doi: 10.1080/13548506.2011.554645] [Medline: 21749242]
2. Sandfort TGM, Bakker F, Schellevis FG, Vanwesenbeeck I. Sexual orientation and mental and physical health status: findings from a Dutch population survey. Am J Public Health 2006 Jun;96(6):1119-1125. [doi: 10.2105/AJPH.2004.058891] [Medline: 16670235]
3. Salway T, Ross LE, Fehr CP, Burley J, Asadi S, Hawkins B, et al. A systematic review and meta-analysis of disparities in the prevalence of suicide ideation and attempt among bisexual populations. Arch Sex Behav 2019 Jan;48(1):89-111. [doi: 10.1007/s10508-018-1150-6] [Medline: 29492768]

XSL•FO
RenderX

4. Bostwick WB, Boyd CJ, Hughes TL, McCabe SE. Dimensions of sexual orientation and the prevalence of mood and anxiety disorders in the United States. Am J Public Health 2010 Mar;100(3):468-475 [FREE Full text] [doi: 10.2105/AJPH.2008.152942] [Medline: 19696380]
5. Alvy LM, McKirnan DJ, Mansergh G, Koblin B, Colfax GN, Flores SA, Project MIX Study Group. Depression is associated with sexual risk among men who have sex with men, but is mediated by cognitive escape and self-efficacy. AIDS Behav 2011 Aug;15(6):1171-1179. [doi: 10.1007/s10461-010-9678-z] [Medline: 20217471]
6. Nouri E, Moradi Y, Moradi G. What is the global prevalence of depression among men who have sex with men? A systematic review and meta-analysis. Ann Gen Psychiatry 2022 Sep 12;21(1):38 [FREE Full text] [doi: 10.1186/s12991-022-00414-1] [Medline: 36096814]
7. King M, Semlyen J, Tai SS, Killaspy H, Osborn D, Popelyuk D, et al. A systematic review of mental disorder, suicide, and deliberate self harm in lesbian, gay and bisexual people. BMC Psychiatry 2008 Aug 18;8:70 [FREE Full text] [doi: 10.1186/1471-244X-8-70] [Medline: 18706118]
8. Meyer IH. Prejudice, social stress, and mental health in lesbian, gay, and bisexual populations: conceptual issues and research evidence. Psychol Bull 2003 Sep;129(5):674-697 [FREE Full text] [doi: 10.1037/0033-2909.129.5.674] [Medline: 12956539]
9. Byrne J, Russel C, Quan A. Legal gender recognition: a multi-country legal and policy review in Asia. Asia Pacific Transgender Network and United Nations Development Programme. 2017. URL: https://weareaptn.org/wp-content/uploads/2019/05/UPDATED-ACKNOWLEDGMENTS-2019_rbap-hhd-2017-legal-gender-recognition-1.pdf [accessed 2024-03-05]
10. The constitution of Nepal. National Law Commission: Nepal. 2015. URL: https://ag.gov.np/files/Constitution-of-Nepal_2072_Eng_www.moljpa.gov_.npDate-72_11_16.pdf [accessed 2024-03-05]
11. UNDP, USAID. Being LGBT in Asia: Nepal country report. United Nations Development Programme and United States Agency for International Development. 2014. URL: https://www.undp.org/asia-pacific/publications/being-lgbt-asia-nepal-country-report [accessed 2024-03-05]
12. Storm S, Deuba K, Shrestha R, Pandey LR, Dahal D, Shrestha MK, et al. Social and structural factors associated with depression and suicidality among men who have sex with men and transgender women in Nepal. BMC Psychiatry 2021 Sep 29;21(1):476 [FREE Full text] [doi: 10.1186/s12888-021-03477-8] [Medline: 34587934]
13. Paudel K, Gautam K, Bhandari P, Wickersham JA, Dhakal M, Sharma S, et al. Suicidal ideation, plan, and attempt among men who have sex with men in Nepal: Findings from a cross-sectional study. PLOS Glob Public Health 2023;3(11):e0002348 [FREE Full text] [doi: 10.1371/journal.pgph.0002348] [Medline: 37992034]
14. Giri PD, Adhikari A, Pradhan M, Yogi I, Khanal S. Barriers in access to health care services among lesbian, gay, bisexual, transgender (LGBT). Int J New Technol Res 2019;5(3):15-20. [doi: 10.31871/IJNTR.5.3.24]
15. Oli N, Onta SR. Self-perception of stigma and discrimination among men having sex with men. J Nepal Health Res Counc 2012 Sep;10(22):197-200. [Medline: 23281450]
16. Singh S, Pant SB, Dhakal S, Pokhrel S, Mullany LC. Human rights violations among sexual and gender minorities in Kathmandu, Nepal: a qualitative investigation. BMC Int Health Hum Rights 2012 May 16;12:7 [FREE Full text] [doi: 10.1186/1472-698X-12-7] [Medline: 22591775]
17. Subedi P, Jha A. Status of healthcare in LGBTQI+ community in Nepal: challenges and possibilities. JNMA J Nepal Med Assoc 2023 Jan 01;61(257):95-97 [FREE Full text] [doi: 10.31729/jnma.7948] [Medline: 37203914]
18. Rauschenberg C, Schick A, Hirjak D, Seidler A, Paetzold I, Apfelbacher C, et al. Evidence synthesis of digital interventions to mitigate the negative impact of the COVID-19 pandemic on public mental health: rapid meta-review. J Med Internet Res 2021 Mar 10;23(3):e23365 [FREE Full text] [doi: 10.2196/23365] [Medline: 33606657]
19. Gilbey D, Morgan H, Lin A, Perry Y. Effectiveness, acceptability, and feasibility of digital health interventions for LGBTIQ+ young people: systematic review. J Med Internet Res 2020 Dec 03;22(12):e20158 [FREE Full text] [doi: 10.2196/20158] [Medline: 33270039]
20. Ministry of Health and Population, New ERA, The DHS Program ICF. Nepal demographic and health survey 2022: key indicators report. Ministry of Health and Population. Kathmandu, Nepal; 2022. URL: https://dhsprogram.com/pubs/pdf/PR142/PR142.pdf [accessed 2024-03-05]
21. Sherin KM, Sinacore JM, Li XQ, Zitter RE, Shakil A. HITS: a short domestic violence screening tool for use in a family practice setting. Fam Med 1998;30(7):508-512. [Medline: 9669164]
22. Kroenke K, Spitzer RL, Williams JBW. The PHQ-9: validity of a brief depression severity measure. J Gen Intern Med 2001 Sep;16(9):606-613 [FREE Full text] [doi: 10.1046/j.1525-1497.2001.016009606.x] [Medline: 11556941]
23. Kim JE, Saw A, Zane N. The influence of psychological symptoms on mental health literacy of college students. Am J Orthopsychiatry 2015 Nov;85(6):620-630 [FREE Full text] [doi: 10.1037/ort0000074] [Medline: 26052815]
24. Klineberg E, Biddle L, Donovan J, Gunnell D. Symptom recognition and help seeking for depression in young adults: a vignette study. Soc Psychiatry Psychiatr Epidemiol 2011 Jun;46(6):495-505. [doi: 10.1007/s00127-010-0214-2] [Medline: 20358174]
25. Swami V. Mental health literacy of depression: gender differences and attitudinal antecedents in a representative British sample. PLoS One 2012;7(11):e49779 [FREE Full text] [doi: 10.1371/journal.pone.0049779] [Medline: 23166769]

XSL•FO
RenderX

26. Garcia C, Grant E, Treharne GJ, Arahanga-Doyle H, Lucassen MFG, Scarf D, et al. 'Is it worth potentially dealing with someone who won't get it?': LGBTQA+ university students' perspectives on mental health care. J R Soc New Zealand 2023 Jul 16:1-15 [FREE Full text] [doi: 10.1080/03036758.2023.2235297]
27. Narendorf SC, Munson MR, Ben-David S, Cole AR, Scott LD. Race and gender differences in attitudes toward help seeking among marginalized young adults with mood disorders: a mixed-methods investigation. Psychiatr Rehabil J 2018 Dec;41(4):277-289. [doi: 10.1037/prj0000312] [Medline: 30507242]
28. Shrestha R, Altice FL, Khati A, Azwa I, Gautam K, Gupta S, et al. Clinic-integrated smartphone app (JomPrEP) to improve uptake of HIV testing and pre-exposure prophylaxis among men who have sex with men in Malaysia: mixed methods evaluation of usability and acceptability. JMIR Mhealth Uhealth 2023 Feb 16;11:e44468 [FREE Full text] [doi: 10.2196/44468] [Medline: 36795465]
29. Drydakis N. mHealth apps and physical and mental health outcomes of sexual minorities. J Homosex 2023 Dec 06;70(14):3421-3448. [doi: 10.1080/00918369.2022.2095240] [Medline: 35904851]
30. Dubov A, Fraenkel L, Goldstein Z, Arroyo H, McKellar D, Shoptaw S. Development of a smartphone app to predict and improve the rates of suicidal ideation among transgender persons (TransLife): qualitative study. J Med Internet Res 2021 Mar 24;23(3):e24023 [FREE Full text] [doi: 10.2196/24023] [Medline: 33596181]
31. Hsieh N, Shuster SM. Health and health care of sexual and gender minorities. J Health Soc Behav 2021 Sep;62(3):318-333 [FREE Full text] [doi: 10.1177/00221465211016436] [Medline: 34528481]
32. Cronin TJ, Pepping CA, Halford WK, Lyons A. Mental health help-seeking and barriers to service access among lesbian, gay, and bisexual Australians. Australian Psychologist 2021 Mar 17;56(1):46-60. [doi: 10.1080/00050067.2021.1890981]
33. Dunbar MS, Sontag-Padilla L, Ramchand R, Seelam R, Stein BD. Mental health service utilization among lesbian, gay, bisexual, and questioning or queer college students. J Adolesc Health 2017 Sep;61(3):294-301. [doi: 10.1016/j.jadohealth.2017.03.008] [Medline: 28549595]
34. Samarasekera U. The rise of racial minority health apps. Lancet Digit Health 2022 Apr;4(4):e218-e219 [FREE Full text] [doi: 10.1016/S2589-7500(22)00043-7] [Medline: 35337641]
35. Beattie A, Shaw A, Kaur S, Kessler D. Primary-care patients' expectations and experiences of online cognitive behavioural therapy for depression: a qualitative study. Health Expect 2009 Mar;12(1):45-59 [FREE Full text] [doi: 10.1111/j.1369-7625.2008.00531.x] [Medline: 19250152]
36. Hatzenbuehler ML, Pachankis JE. Stigma and minority stress as social determinants of health among lesbian, gay, bisexual, and transgender youth: research evidence and clinical implications. Pediatr Clin North Am 2016 Dec;63(6):985-997. [doi: 10.1016/j.pcl.2016.07.003] [Medline: 27865340]
37. Chard AN, Metheny NS, Sullivan PS, Stephenson R. Social stressors and intoxicated sex among an online sample of men who have sex with men (MSM) drawn from seven countries. Subst Use Misuse 2018 Jan 02;53(1):42-50. [doi: 10.1080/10826084.2017.1322985] [Medline: 28792280]
38. Dowshen N, Binns HJ, Garofalo R. Experiences of HIV-related stigma among young men who have sex with men. AIDS Patient Care STDS 2009 May;23(5):371-376 [FREE Full text] [doi: 10.1089/apc.2008.0256] [Medline: 19320600]
39. Khati A, Wickersham JA, Rosen AO, Luces JRB, Copenhaver N, Jeri-Wahrhaftig A, et al. Ethical issues in the use of smartphone apps for HIV prevention in Malaysia: focus group study with men who have sex with men. JMIR Form Res 2022;6(12):e42939 [FREE Full text] [doi: 10.2196/42939] [Medline: 36563046]
40. Shrestha R, Fisher C, Wickersham JA, Khati A, Kim R, Azwa I, et al. Privacy and confidentiality concerns related to the use of mHealth apps for HIV prevention efforts among Malaysian men who have sex with men: cross-sectional survey study. JMIR Form Res 2021 Dec 16;5(12):e28311 [FREE Full text] [doi: 10.2196/28311] [Medline: 34924355]
41. Martínez-Pérez B, de la Torre-Díez I, López-Coronado M. Privacy and security in mobile health apps: a review and recommendations. J Med Syst 2015 Jan;39(1):181. [doi: 10.1007/s10916-014-0181-3] [Medline: 25486895]
42. Morera EP, de la Torre Díez I, Garcia-Zapirain B, López-Coronado M, Arambarri J. Security recommendations for mHealth apps: elaboration of a developer's guide. J Med Syst 2016 Jun;40(6):152. [doi: 10.1007/s10916-016-0513-6] [Medline: 27147515]
43. Rajan S, Rathod SD, Luitel NP, Murphy A, Roberts T, Jordans MJD. Healthcare utilization and out-of-pocket expenditures associated with depression in adults: a cross-sectional analysis in Nepal. BMC Health Serv Res 2020 Mar 25;20(1):250 [FREE Full text] [doi: 10.1186/s12913-020-05094-9] [Medline: 32213188]
44. Regmi SK, Pokharel A, Ojha SP, Pradhan SN, Chapagain G. Nepal mental health country profile. Int Rev Psychiatry 2004;16(1-2):142-149. [doi: 10.1080/09540260310001635186] [Medline: 15276946]
45. Knapp M, Funk M, Curran C, Prince M, Grigg M, McDaid D. Economic barriers to better mental health practice and policy. Health Policy Plan 2006 May;21(3):157-170 [FREE Full text] [doi: 10.1093/heapol/czl003] [Medline: 16522714]
46. Charlton BM, Gordon AR, Reisner SL, Sarda V, Samnaliev M, Austin SB. Sexual orientation-related disparities in employment, health insurance, healthcare access and health-related quality of life: a cohort study of US male and female adolescents and young adults. BMJ Open 2018 Jul 26;8(6):e020418 [FREE Full text] [doi: 10.1136/bmjopen-2017-020418] [Medline: 30049672]

XSL•FO
RenderX

47. Meiksin R, Melendez-Torres GJ, Miners A, Falconer J, Witzel TC, Weatherburn P, et al. E-Health interventions targeting STIs, sexual risk, substance use and mental health among men who have sex with men: four systematic reviews. Public Health Res 2022;10(4):1-322 [FREE Full text] [doi: 10.3310/BRWR6308] [Medline: 35344312]
48. Lucassen MFG, Hatcher S, Fleming TM, Stasiak K, Shepherd MJ, Merry SN. A qualitative study of sexual minority young people's experiences of computerised therapy for depression. Australas Psychiatry 2015 Jun;23(3):268-273. [doi: 10.1177/1039856215579542] [Medline: 25881962]
49. Anthony P, Crawford P. Service user involvement in care planning: the mental health nurse's perspective. J Psychiatr Ment Health Nurs 2000 Oct;7(5):425-434. [doi: 10.1046/j.1365-2850.2000.00332.x] [Medline: 11933469]
50. McNair RP, Hegarty K. Guidelines for the primary care of lesbian, gay, and bisexual people: a systematic review. Ann Fam Med 2010;8(6):533-541 [FREE Full text] [doi: 10.1370/afm.1173] [Medline: 21060124]

Abbreviations

FG: focus group
mHealth: mobile health
MSM: men who have sex with men
LGBTIQA+: lesbian, gay, bisexual, transgender, intersex, queer, and asexual

Edited by G Eysenbach, A Kushniruk; submitted 02.01.24; peer-reviewed by A AL-Asadi, A Nyamathi; comments to author 31.01.24; revised version received 01.02.24; accepted 18.02.24; published 29.03.24.

Please cite as:
Gautam K, Aguilar C, Paudel K, Dhakal M, Wickersham JA, Acharya B, Sapkota S, Deuba K, Shrestha R
Preferences for mHealth Intervention to Address Mental Health Challenges Among Men Who Have Sex With Men in Nepal: Qualitative Study
JMIR Hum Factors 2024;11:e56002
URL: https://humanfactors.jmir.org/2024/1/e56002
doi:10.2196/56002
PMID:38551632

©Kamal Gautam, Camille Aguilar, Kiran Paudel, Manisha Dhakal, Jeffrey A Wickersham, Bibhav Acharya, Sabitri Sapkota, Keshab Deuba, Roman Shrestha. Originally published in JMIR Human Factors (https://humanfactors.jmir.org), 29.03.2024. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on https://humanfactors.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Patient Perspectives on Communication Pathways After Orthopedic Surgery and Discharge and Evaluation of Team-Based Digital Communication: Qualitative Exploratory Study

Lili Worre Høpfner Jensen[1], BSc, MSc; Ole Rahbek[1], MD, PhD; Rikke Emilie Kildahl Lauritsen[1], BSc, MSc; Søren Kold[1], MD, PhD; Birthe Dinesen[2], MSc, PhD

[1]Interdisciplinary Orthopaedics, Orthopaedic Surgery Department, Aalborg University Hospital, Aalborg, Denmark

[2]Laboratory for Welfare Technologies – Digital Health & Rehabilitation, Sport Sciences – Performance & Technology, Department of Health Science and Technology, Aalborg University, Aalborg East, Denmark

Corresponding Author:
Lili Worre Høpfner Jensen, BSc, MSc
Interdisciplinary Orthopaedics
Orthopaedic Surgery Department
Aalborg University Hospital
Hobrovej 18-22
Aalborg, 9000
Denmark
Phone: 45 60229406
Email: lili.jensen@rn.dk

Abstract

Background: The transition from hospital to home after orthopedic surgery requires smooth communication and coordination between patients and their team of care to avoid fragmented care pathways. Digital communication is increasingly being used to facilitate easy and accessible asynchronous communication between patients and health care professionals across settings. A team-based approach to digital communication may provide optimized quality of care in the postoperative period following orthopedic surgery and hospital discharge.

Objective: This study was divided into two phases that aimed to (1) explore the perspectives of patients undergoing orthopedic surgery on current communication pathways at a tertiary hospital in Denmark and (2) test and explore patients' experiences and use of team-based digital communication following hospital discharge (eDialogue).

Methods: A triangulation of qualitative data collection techniques was applied: document analysis, participant observations (n=16 hours), semistructured interviews with patients before (n=31) and after (n=24) their access to eDialogue, and exploration of use data.

Results: Findings show that patients experience difficult communication pathways after hospital discharge and a lack of information due to inadequate coordination of care. eDialogue was used by 84% (26/31) of the patients, and they suggested that it provided a sense of security, coherence, and proximity in the aftercare rearranging communication pathways for the better. Specific drivers and barriers to use were identified, and these call for further exploration of eDialogue.

Conclusions: In conclusion, patients evaluated eDialogue positively and suggested that it could support them after returning home following orthopedic surgery.

(JMIR Hum Factors 2024;11:e49696) doi:10.2196/49696

KEYWORDS

digital communication; patient-provider communication; continuity of care; interdisciplinary communication; hospital discharge; orthopedic surgery; postoperative care; text messaging; mobile phone

XSL•FO
RenderX

Introduction

Across the health care system, digital communication is being implemented as an addition to traditional communication pathways [1,2]. Digital communication is a form of eHealth [3] that facilitates asynchronous 2-way text messaging between patients and health care professionals (HCPs). Digital communication is typically facilitated through email [4,5]; secure text messaging in patient portals [2,6]; or as a feature in mobile health apps developed for specific purposes, for example, postoperative monitoring [7,8] and neonatal tele-homecare [9,10]. Establishing the effects of using digital communication is still challenging [11,12]; however, an increasing number of studies suggest that it can support patients in taking care of their own health [12] and address unmet communication needs after hospital discharge [13,14]. When digital communication is used with the purpose of facilitating team-based communication across settings, studies indicate that it may contribute to improving continuity of care (COC) in transitions from hospital to home [14-16]. COC is essential for patients undergoing complex and long-term procedures [17]. Patients who receive care across time and settings are susceptible to fragmented care, and the absence of consistent professional support and communication may lead to neglect that ultimately affects patient safety [18-21]. Because of the growing population in need of orthopedic surgery, workforce shortage [22], and optimized surgery techniques, patients undergoing orthopedic surgery are discharged earlier [23]. Day surgery is increasingly used, and even patients undergoing complex treatments are hospitalized for a shorter time. Common to patients undergoing orthopedic surgery is a need for continuing rehabilitation across settings, supported by adequate communication and home symptom monitoring between follow-up visits [24,25]. Even so, only a few studies have addressed the use of team-based digital communication involving patients and HCPs across settings, and primarily in other patient populations, such as patients with cancer [14,15,26] and children with cerebral palsy [27]. To our knowledge, no studies have investigated the use of team-based digital communication after hospital discharge in orthopedic surgery, although these patients often have long periods of rehabilitation, where cross-disciplinary and cross-sectoral communication is pivotal [28].

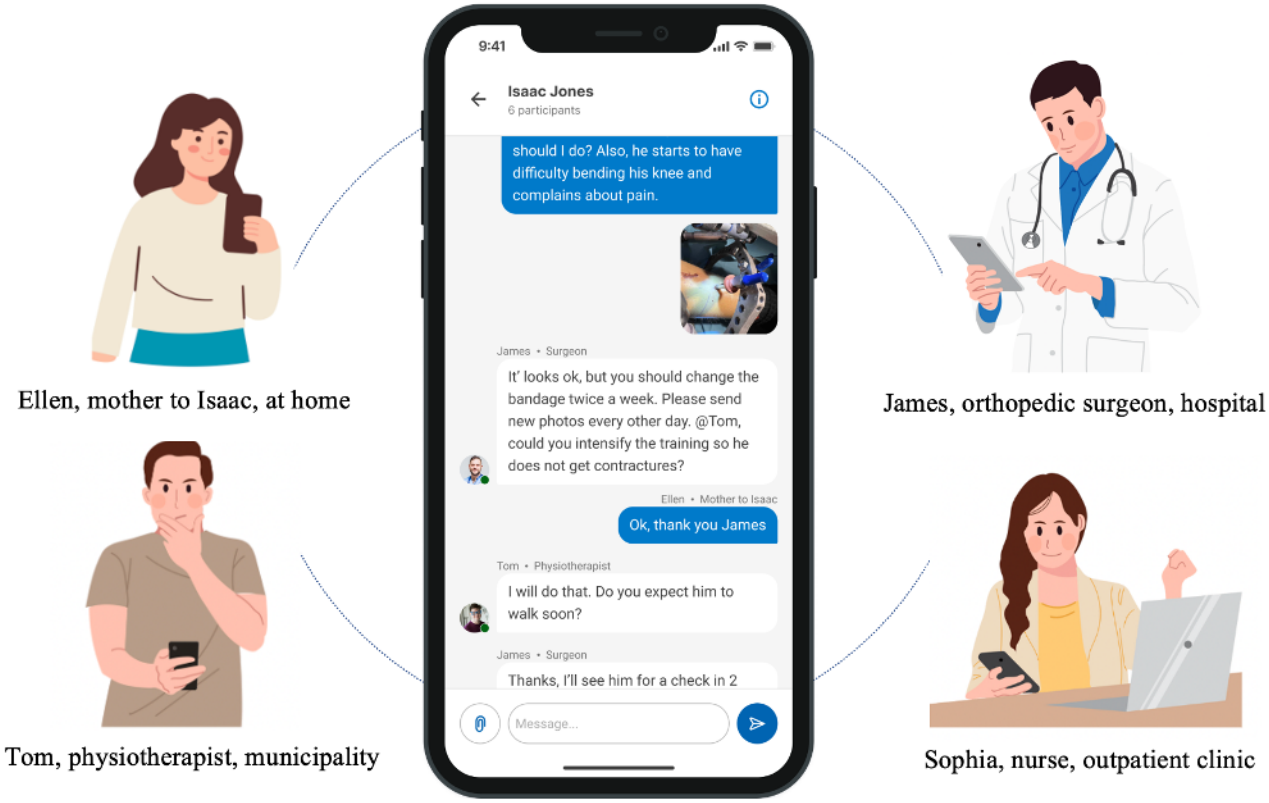
Therefore, the aim of this study was to explore the perspectives of patients undergoing orthopedic surgery on current communication pathways (phase 1) and to subsequently test and explore their experiences and use of a team-based digital communication solution (eDialogue) to evaluate whether the solution can support their needs after hospital discharge (phase 2).

Methods

The eDialogue Intervention

The technical solution used in this study was a simple General Data Protection Regulation–compliant solution, developed for team-based communication, that lets users chat directly with each other with texts and photos (“LetDialog” by Visma) [29] (Figure 1).

Figure 1. Illustration of the team-based digital communication (eDialogue) used in this study, where patients and health care professionals across settings could text and send photos to communicate about postdischarge issues.



XSL•FO
RenderX

The solution was accessed through an app for smartphones or through a website. Users could choose how they accessed it individually. To ensure compliance with the current legislation, user profiles were created with a digital signature (NemID), and the digital dialogues were stored in a secure cloud-based solution. A data processor agreement was made among the North Denmark Region, Aalborg University Hospital, and Visma before this study.

The features were basic asynchronous text messaging and exchange of photos. Photos could be taken directly or uploaded and sent through the solution for review by the health care team. Team-based digital communication was organized in teams, defined by the individual patient, in a shared chat. Notifications were sent to all the participants when there were new posts. Key HCPs from the orthopedic surgery department at the hospital were identified and recruited for participation before the study (surgeons, nurses, and physiotherapists). Other HCPs from municipal or private settings were recruited ad hoc and based on patients' wishes (eg, physiotherapists from the municipality).

Study Design

The study was exploratory, using a triangulation of qualitative data collection techniques, including document analysis, participant observations [30], semistructured interviews [31], and use data, with the purpose of obtaining in-depth knowledge of patients' perspectives and the context.

Theoretical Framework

The theoretical framework for this study was inspired by the concepts of COC [17,32], which is used as a measure of quality of care in health care transitions. COC includes *informational continuity*, described as the use of medical or personal information to provide appropriate care over time; *management continuity*, which refers to the provision of timely, coordinated, and complementary services that are responsive to patients' needs to connect care over time; and *relational continuity*, which involves the consistency and quality of relationships between patients and providers as a means of connecting care over time [32]. All 3 dimensions should be integrated to achieve COC, and thus, COC is maximized when planning for patient-provider continuity, information exchange, and seamless coordination of services in the period of transition from hospital to home [32-34]. For this study, COC has inspired the data collection and analysis of interviews and observations as well as the use of team-based digital communication to prevent fragmented care experiences after hospital discharge.

Participants and Setting

The study was conducted at the Orthopedic Surgery Department of Aalborg University Hospital, Denmark. The recruitment of participants began in May 2021 and ended in November 2021. The final follow-up interviews were conducted 2 months later in January 2022.

In phase 1, participants were recruited consecutively based on predefined inclusion criteria: (1) patients, or their parents if the patient was aged <15 years, undergoing deformity correction (DC) surgery or anterior cruciate ligament (ACL) reconstruction; (2) those who were able to read and write Danish; (3) those who were discharged to their own home and had planned follow-up in the outpatient clinic; and (4) those who owned a smartphone and had access to a secure digital signature. The exclusion criteria were (1) those who were not able to understand Danish and (2) those who were not cognitively able to participate in interviews.

The 2 patient groups, DC and ACL, were selected because they represent 2 different orthopedic surgical care pathways. Involving both patient groups allowed us to gain an insight into the different needs of patients undergoing orthopedic surgery. ACL is performed as a day surgery (ie, discharge on the same day), whereas patients in the DC group most often have longer hospitalizations and prolonged treatments.

The same recruitment procedure was used for patients undergoing DC or ACL. The patients were approached by secretaries at the hospital with an invitation to participate. If the patients agreed to be called by phone with information about participation in the study, the first author (LWHJ) would call them to provide oral participant information. Written participant information was then sent by email, and the patients were given time to consider participation. One patient did not want to participate after receiving oral information due to a lack of mental capacity to participate in the interviews. Another patient could not be contacted by telephone after he had initially registered his telephone number. Both patients were from the ACL group.

In phase 2, patients and parents (if the patient was a minor) were onboarded to eDialogue on the day of discharge. The orthopedic surgeon, who had performed the surgery, was invited to join the patients' dialogue, as were nurses from the outpatient clinic and physiotherapists across sectors who were involved in the patient's care and rehabilitation after discharge. Thus, the patients were connected with known HCPs and were able to use eDialogue as needed from the day of discharge until 2 months after discharge. The patients could send texts and photos whenever it suited them, but they were told that a 24-hour response time on weekdays (Monday to Friday) would be aimed for. As such, messages sent during weekends and holidays would be responded to on the next weekday. It was pointed out, both verbally and in the participant information letter, that in case of emergency, patients should not use the solution but instead call, as they usually would have done before access to eDialogue. Thus, eDialogue was an addition to traditional communication channels (eg, telephone calls and email) and an extra opportunity for communication after discharge.

Data Collection

A triangulation of data collection techniques was performed to achieve exhaustive knowledge of current communication pathways, patients' perspectives, and their experiences with eDialogue.

Phase 1

First, document analysis was performed on documents and guidelines for postdischarge communication between patients and HCPs followed by participant observations of workflows (n=16 hours). The aim of the document analysis was to obtain knowledge of the policies and context of the study. The aim of

XSL•FO
RenderX

observations was to document the current communication pathways for patients following hospital discharge. Participant observations were performed by LWHJ and followed a predefined observation guide [30]. Observations were carried out at the orthopedic surgery ward and the outpatient clinic at the hospital and documented in Word files (Microsoft Corp). This involved, for example, secretaries' handling of incoming phone calls from patients, registration of patient inquiries, procedures for passing on messages to nurses and orthopedic surgeons, and HCPs' calls with patients. In addition, existing systems for communication with discharged patients were reviewed, including written communication to patients via "E-box," (a secure digital mail system for communication from Danish authorities) correspondence between HCPs across hospitals and municipalities in the local electronic health record, and interprofessional communication related to patients' phone calls.

Second, semistructured interviews were conducted at the point of inclusion for each participant (N=31). The aim was to explore patients' and parents' perspectives on current communication pathways. Interviews were performed using video 5 to 7 days before surgery for patients from the ACL group (n=14) and physically at the ward for patients and parents from the DC group (n=17) because they were all hospitalized in connection with their operation. All interviews were conducted by LWHJ based on a predefined semistructured interview guide (Multimedia Appendix 1). The guide was developed based on the theoretical framework for this study and combined with exploratory questions. It was pilot-tested in 2 patients similar to the study participants and revised accordingly. The interviews were carried out until data saturation had been reached, defined by the point where no new insights into participants' responses occurred, indicating the achievement of a comprehensive understanding of the participants' perspectives [31]. The interviews were audio recorded using a digital voice recorder (DM-450; Olympus) and lasted for 40 to 60 minutes. They were continuously transcribed and documented in Word files. During and at the end of each interview, key points were summarized to ensure the credibility of the meanings expressed.

Phase 2

Semistructured follow-up interviews were performed with the same patients and parents 2 months after hospital discharge (24/31, 77%). The aim was to explore their experiences of using eDialogue for team-based communication in the postdischarge period. The interviews were performed by LWHJ, audio recorded, and followed a predefined interview guide that was pilot-tested (Multimedia Appendix 1). The interviews were conducted until data saturation was reached for each patient group [31]. They lasted between 30 and 60 minutes. Both users and those who did not use eDialogue after getting access were interviewed. A total of 6 patients (DC: n=3; ACL: n=3) were reached by phone, their experiences were discussed, and a short report was written. Nothing new emerged from these conversations. One parent of a child from the DC group was lost to follow-up as she did not return our calls. Interviews were performed face-to-face at the ward or digitally based on the preferences of the participants. Participants were most likely to choose web-based interviews due to convenience and distance to the hospital, and data collection was conducted at the same time as the COVID-19 pandemic.

Use data of eDialogue was collected through registration of events and manual counts of messages exchanged in all digital dialogues. Data included the total number of messages exchanged in eDialogue during the 2-month study period, the number of text messages and photos sent by patients or parents, and the number of text messages that actually needed a reply from HCPs. In addition, the distribution of text messages per week per patient group was collected and displayed to show the differences between groups. Content analysis [31] of the messages sent by the patients and the parents was performed to provide insight into question categories as well as how they were distributed between the patient groups.

Data Analysis

Data analysis was carried out in NVivo (version 20.6.2; Lumivero), inspired by Brinkmann and Kvale [31], with the aim of achieving an in-depth understanding and connection of the participants' expressed perspectives on current communication pathways (phase 1) and experiences using eDialogue (phase 2).

Separate data analyses were carried out for phase 1 and phase 2 and for each patient group (DC and ACL), all involving 3 steps: meaning coding, meaning condensation, and meaning interpretation (Textbox 1).

In phase 1, observational data were integrated into the data set to enhance the understanding of existing communication pathways for patients in need of postdischarge contact.

Use data from eDialogue were analyzed and presented using simple descriptive statistics and basic content analysis to present the overall question categories.

The reporting of this study followed the Consolidated Criteria for Reporting Qualitative Research checklist [35].

XSL•FO
RenderX

Textbox 1. The process of the thematic analysis.

Meaning coding

- Coding of the transcribed interviews was initially performed individually by 2 of the authors (LWHJ and REKL) by randomly selecting 4 interviews from each patient group and from each phase. Coding was conducted with an inductive approach with the aim of reflecting the meanings expressed by the participants.
- To achieve intersubjectivity before and during the analysis, interviews were individually read and reread by LWHJ and REKL, notes were made for initial ideas for codes, and these were then compared and discussed until agreement. In phase 1, this resulted in 21 codes in the deformity correction (DC) group and 18 codes in the anterior cruciate ligament (ACL) group. In phase 2, we identified 18 codes in the DC group and 15 codes in the ACL group. LWHJ continued the data analysis of the remaining interviews by applying the same codes to the entire data.

Meaning condensation

- The codes were then reread and condensed in discussion with REKL and the last author (BD) through several iterations, and this process resulted in 12 codes for the DC group and 8 codes for the ACL group in phase 1 and 11 codes for the DC group and 7 codes for the ACL group in phase 2. We merged and we left behind codes that did not directly address the research questions or were only described vaguely by 1 participant.
- The remaining codes were then discussed with all authors to achieve further condensation and to define and name subthemes and themes that would capture the essence of the data. It was clear to the authors that the 2 patient groups had expressed similar perspectives on the phenomena of interest, and therefore, codes could be merged between the 2 groups in this step. Theme generation was based on a systematic identification and organization of recurring patterns, topics, or concepts within the data set. This process resulted in 3 overall themes and 6 subthemes for phase 1 and 3 overall themes and 3 subthemes for phase 2.

Meaning interpretation

- Themes were defined and described narratively, and data extracts were chosen for presentation in the manuscript before writing the findings.

Ethical Considerations

Before the study started, the Ethics Committee of Northern Jutland was approached, and it was found that the study did not require approval, as eDialogue was an extra opportunity for patients to communicate directly with their team of HCPs across sectors. This was confirmed by email on March 18, 2021 (2021-000438). The study was registered with the Regional Committee on Health Research and approved (ID number 2021-057). All participants received thorough oral and written information and guidance in the use of eDialogue before discharge. The study followed the Helsinki Declaration, and the participants signed an informed consent form and were able to leave the study without explanation or effects on usual care. All patients or parents had access to eDialogue for 2 months after hospital discharge. If they wanted, patients were allowed to keep the possibility of eDialogue with their team of HCPs after 2 months and until their follow-up in the outpatient clinic was completed. An administrator from the project group was passively present in all dialogues to continuously observe whether the patients used the solution for emergencies against the given advice.

Results

Participants' Characteristics

Table 1 provides the baseline description of the 31 patients included in this study. The patients were recruited from 2 different subgroups of orthopedic surgery: DC (17/31, 55%) and ACL (14/31, 45%).

The patients in the DC group were, for example, patients with malalignment or limb length discrepancy, and they were all hospitalized for >1 day. The patients in the ACL group were all treated with ACL reconstruction, and they had the procedure performed as day surgery. Of the 14 patients with ACL injuries, 7 (50%) had a concurrent meniscal injury.

Across the groups, most patients were male (22/31, 71%), and patients ranged in age from 1 to 59 years. Patients from the DC group were discharged from the hospital after an average of 6.1 (range 1-9) days, and patients from the ACL group were all discharged on the same day of surgery (<9 hours of admission). All the included patients were discharged to their own home. Of the 31 patients included, 14 (45%) had previously undergone orthopedic surgery at Aalborg University Hospital, and thus, they were able to reflect on previous experiences with postdischarge communication during the initial interviews in phase 1. In the DC group, 5 patients lived outside the North Jutland Region.

A total of 42% (13/31) of the patients were children aged <15 years, and thus, their parents were the primary users of eDialogue. Therefore, the baseline characteristics of all users of eDialogue are presented in Table 2.

The table shows 33 users in total because 2 patients aged 16 and 17 years had a parent joining the dialogue with them. Of the 13 parents who were users of eDialogue with or on behalf of their child, 77% (10/13) were female (mothers). The mean age of the parents was 43 (range 37-48) years on the day of discharge. All users of eDialogue used a smartphone on a daily basis.

XSL•FO
RenderX

Table 1. Characteristics of all patients across groups (DC[a], n=17; ACL[b], n=14; N=31).

Characteristics	Values
Sex (DC/ACL), n (%)	
Female	5 (29)/4 (29)
Male	12 (71)/10 (71)
Age at discharge (years), mean (range)	
DC	19.2 (1-59)
ACL	29.1 (17-46)
Length of hospital stay, mean (range)	
DC	6.1 (1-9) days
ACL	1 (7-9) hours
Previously had orthopedic surgery (yes/no), n (%)	
DC	12 (71)/5 (29)
ACL	2 (14)/12 (86)
Highest education level (DC/ACL), n (%)	
Primary or high school	12 (71)/5 (36)
Vocational education (skilled worker)	2 (12)/2 (14)
Short education, 2-3 years	1 (6)/2 (14)
Bachelor's degree, 3-5 years	2 (12)/4 (29)
Academic education, 5-8 years	0 (0)/1 (7)
Work status (DC/ACL), n (%)	
Student	13 (76)/7 (50)
Unemployed	1 (6)/2 (14)
Employed	3 (18)/5 (36)
Civil status (DC/ACL), n (%)	
Living alone	3 (18)/4 (29)
Cohabiting	14 (82)/10 (71)

[a]DC: deformity correction.
[b]ACL: anterior cruciate ligament.

XSL•FO
RenderX

Table 2. Baseline characteristics of all users of eDialogue (DC[a], n=18; ACL[b], n=15; patients and parents; N=33).

Characteristics	Values
Distribution of users (DC/ACL), n (%)	
Patients	6 (33)/14 (93)
Parents	12 (67)/1 (7)
Sex (DC/ACL), n (%)	
Female	12 (67)/5 (33)
Male	6 (33)/10 (67)
Age at discharge (years), mean (range)	
DC	39.8 (16-59)
ACL	28.8 (17-46)
Highest education level (DC/ACL), n (%)	
Primary or high school	2 (11)/5 (33)
Vocational education (skilled worker)	3 (17)/2 (13)
Short education, 2-3 years	3 (17)/2 (13)
Bachelor's degree, 3-5 years	8 (44)/5 (33)
Master's degree, 5-8 years	2(11)/1 (7)
Work status (DC/ACL), n (%)	
Student	2 (11)/7 (47)
Unemployed	1 (6)/2 (13)
Employed	14 (78)/6 (40)
Disability pensioner	1 (6)/0 (0)
Civil status (DC/ACL), n (%)	
Living alone	4 (22)/3 (20)
Cohabiting	14 (78)/12 (80)

[a]DC: deformity correction.

[b]ACL: anterior cruciate ligament.

Phase 1: Perspectives on Current Communication Pathways

Themes and Subthemes

Through the initial interviews, 3 themes and associated subthemes were revealed across the groups. Overall, patients and parents from the DC and ACL groups had similar experiences of, and perspectives on, current communication pathways. However, some subthemes were more prominent in one group than the other. This is illustrated by showing how many patients and parents from each group expressed experiences related to the specific subtheme (Table 3).

XSL•FO
RenderX

Table 3. Themes and subthemes of patients' and parents' perspectives on current communication pathways with HCPs[a] after hospital discharge (N=31).

Themes and subthemes	DC[b] (n=17), n (%)	ACL[c] (n=14), n (%)
Difficult communication pathways		
Doubts about who to contact and when	8 (47)	7 (50)
Withhold questions or forget to ask	7 (41)	9 (64)
Lack of information due to inadequate coordination of care		
Knowledge is not shared sufficiently	8 (47)	6 (43)
Hard to be "the messenger" between HCPs	9 (53)	5 (36)
Relations and communication provide "peace of mind"		
Relational continuity matters	15 (88)	4 (29)
Contacts provides a sense of being cared for	10 (59)	2 (14)

[a]HCP: health care professional.

[b]DC: deformity correction.

[c]ACL: anterior cruciate ligament.

Difficult Communication Pathways

Most patients and parents expressed frustrations related to difficult communication pathways when they needed contact with HCPs. They were in doubt about who to contact regarding specific issues both before and after surgery and discharge:

> *It was like a week after discharge, and I didn't know who to ask. Should I contact the department, the outpatient clinic or my own physician? I didn't know that. They kept telling me to call a new location.* [Mother of patient 2, DC]

The patients also described how they would often forget to ask questions at the outpatient clinic or they would withhold questions because they found it difficult to assess whether their issues were "severe enough" to take up HCPs time. A patient explains how it had previously led to concerns and worsening of symptoms:

> *I couldn't lift up my leg like I had been able to before...The next morning, the knee was barely visible due to swelling. Well, I should probably have done something the day before, but I didn't. You just know that when you call the hospital, you must go through several people, and I don't want to be a nuisance either.* [Patient 4, DC]

Lack of Information Due to Inadequate Coordination of Care

Patients in the ACL group highlighted a lack of information before surgery. Similarly, they described missing information in the first weeks after discharge, before their postoperative follow-up visit, and before starting rehabilitation with a physiotherapist:

> *Actually, I didn't know what I was supposed to do. Maybe I didn't ask enough questions before discharge. The first week (after discharge) I didn't do anything. I was wearing this DonJoy bandage and I didn't put stress on my leg or anything. And it turns out that I really should have done that.* [Patient 1, ACL]

They had questions about rehabilitation and restrictions associated with the operation, and this led to Google searches, which usually left them more confused:

> *I felt like I was in a no man's land and didn't really know what to do.* [Patient 3, ACL]

In the DC group, the patients and the parents described how knowledge is not shared across sectors in a sufficient and timely fashion. The fact that HCPs in the municipality did not have specialty-specific knowledge, as did those from the hospital, was perceived as unsafe and uncertain. They described situations in which home care nurses or physiotherapists had little or no experience with their treatment and care. That placed a massive burden on the patients or the parents to be in "control" of everything. Lack of information and coordination across sectors also led to confusion regarding the rehabilitation, for example, when the physiotherapist understood the rehabilitation plan differently than the patient remembered it. The patients and the parents from the DC group pointed out how they become the "messengers" and thus responsible for passing on information between the hospital and municipal providers. They viewed this as burdensome, expressing insecurity about accurately conveying all crucial information:

> *It's the fact that it is our interpretation of what is heard. You know, it is not necessarily medical language that we pass on to the next professional.* [Mother of patient 13, DC]

> *The physiotherapists often ask questions like "what did the surgeon say?" But when you have no professional knowledge, and you are busier with being there for your child, then there might be things I do not remember or consider as being important.* [Mother of patient 12, DC]

Relations and Communication Provide "Peace of Mind"

Patients and parents from both groups highlighted the importance of the relationship and communication with HCPs.

XSL•FO
RenderX

However, they had different perceptions of their actual needs. For the patients in the ACL group, the most important thing was that the HCPs were "competent." This was also valid in the DC group, but they unanimously expressed that the relationship and contact with known HCPs were just as important to them. The mother of a boy, who had been through several operations throughout his childhood, described what the relationship between her son and the HCPs at the hospital meant:

> *It gives, well, it gives you peace. It gives peace of mind even before you have to leave home (to attend surgery or follow-up visit). He can say: "Well, now we're going home to Aalborg again soon," and people will say "You don't live in Aalborg, do you?." And then he would respond: "Well, a lot of my time, I do."* [Mother of patient 7, DC]

The same perspective was elaborated by the mother of another boy:

> *I think it's about safety, trust, and recognizability, and we don't refer to it as the "doctor," we say we're going to see him (the surgeon) or her (the nurse).* [Mother of patient 15, DC]

During the initial interviews, it became clear that some patients undergoing long-term treatments in the DC group already used email or SMS text messaging for communication with the orthopedic surgeon or the physiotherapist. This was described as a workaround because traditional communication pathways did not meet their needs, such as calling the secretary, who would leave a note for the nurse or the surgeon to call the patient. The patients and the parents expressed that it made them feel supported, and thus, they largely understood the intention of eDialogue. When asked about their expectations of eDialogue, most patients and parents who had previous experiences with orthopedic surgery expressed that they wished they had had the opportunity of team-based digital communication the first time. Thus, they expected that their previous experiences of "being a patient" would minimize their need for eDialogue at this time.

Phase 2: Experiences With, and Use of, eDialogue After Discharge

Themes and Subthemes

All 31 patients or their parents included in this study were given access to eDialogue for 2 months after discharge with their team of HCPs across sectors. Interviews with 77% (24/31) of the patients and parents led to 3 overall themes and associated subthemes identified across the groups. As in the initial interviews, some subthemes were more prominent in one group than the other and thus highlighted in the table (Table 4).

Table 4. Themes and subthemes of patients and parents' experiences of using eDialogue with HCPs[a] after discharge (n=24).

Themes and subthemes	DC[b] (n=13), n (%)	ACL[c] (n=11), n (%)
Digitally enhanced coherence and proximity		
A sense of security at home	13 (100)	7 (64)
Sharing knowledge between patients and HCPs	9 (69)	5 (45)
Drivers and barriers to use		
Recognizable, informal tool and easy to use	11 (85)	8 (73)
To "be invited" to dialogue by HCPs allows use	6 (46)	4 (36)
Worry about overburdening HCPs	10 (77)	2 (18)
eDialogue rearranges communication pathways		
Reduces the need for phone calls	12 (92)	6 (55)
Text messages and photos are adequate	9 (69)	7 (64)

[a]HCP: health care professional.

[b]DC: deformity correction.

[c]ACL: anterior cruciate ligament.

Digitally Enhanced Coherence and Proximity

Across groups, patients and parents unanimously reported that the possibility of easy and direct communication with HCPs after discharge provided them with a sense of security at home. Although eDialogue was used sparingly by some patients, the possibility made them feel at ease during the rehabilitation period. For the patients who used eDialogue more, it was expressed that it helped them get through the first period after discharge because they felt "closer" to the HCPs and as if they had a constant "back up":

> *For me, it is very much about security, I almost feel that I have the surgeon by my side all the time. The first time (of surgery and discharge), I felt that he was far away.* [Patient 4, DC]

The patients in the ACL group appreciated the opportunity to ask questions, but the need for communication was most evident in the first weeks after discharge and before the first clinical follow-up and exercise sessions with physiotherapists:

> *Before my first checkup, I encountered some problems that I really wanted answered, so that I didn't have to go and wait and worry if there was something*

XSL•FO
RenderX

> *wrong. It was solved immediately in eDialogue.* [Patient 8, ACL]

For the patients and the parents in the DC group, eDialogue specifically helped HCPs share important information across sectors. They described how no longer being responsible for passing on information between the surgeon and the physiotherapist at the municipality brought relief and was highly appreciated:

> *Then we could see that they had the dialogue and then we knew that when we showed up for training next time, the physiotherapist knew it, so we didn't have to explain, which we found difficult anyway.* [Father of patient 10, DC]

In other cases, the patients described how municipal HCPs would use eDialogue indirectly to keep updated with the patient's progress just by reading the messages exchanged between the HCPs from the hospital and the patient. This provided a basis for a common point of view at the patient's next training session.

The parents of minor children described how they used eDialogue to calm their child or explain the treatment plan to them by reading them messages from HCPs.

Drivers and Barriers to Use

In both groups, the patients and the parents agreed that eDialogue presented as a recognizable and informal tool that was easy to use and that this promoted their use. The short response time was also highlighted as a main reason to use eDialogue:

> *I don't remember a day has passed, more like minutes or hours. So, it's been cool. It would never have been the case if I had to call.* [Patient 1, ACL]

Few patients experienced a late or no response. If it happened with their first question, they explained that it made them lose courage to use eDialogue another time. In general, the patients and the parents felt that the use of eDialogue was less intrusive than calling, but they also expressed worry about overburdening the HCPs. By contrast, they expected HCPs to manage their working hours themselves and assess when they had the time to respond:

> *To begin with, I thought that I would not burden the system unnecessarily...but it probably became a little more urgent and I worried about the way he was feeling, so I texted them and got a reply shortly after.* [Mother of patient 12, DC]

No patients expected answers out of hours, but some sent messages at these times to be relieved. However, they all emphasized that they could have waited for a response until the next weekday. A patient from the ACL group described her reflections about sending a message on a Friday night:

> *And of course, I thought, Oh no, now I hope he doesn't feel obliged to answer, but I also thought that they must be professional and decide for themselves.* [Patient 11, ACL]

Some patients and parents described how, before discharge, some HCPs would urge them to use eDialogue if needed and that the feeling of being invited made them more inclined to use it after coming home. The patients from the ACL group also described how eDialogue opened up the possibility to ask about "minor issues," which they might not have called about.

Among nonusers or those who used eDialogue sparingly, it was expressed that they simply did not have the need, as everything went as planned. Nonuse was also attributed to having frequent follow-ups at the outpatient clinic or attending physiotherapy several times a week.

eDialogue Rearranges Communication Pathways

The patients and the parents highlighted how the use of eDialogue had prevented phone calls or additional physical attendance after discharge; this was particularly prominent for the patients in the DC group:

> *Well, to start with we used eDialogue quite a bit I would say. As soon as we had any questions, we texted them and did not need any other forms of communication.* [Mother of patient 8, DC]

In a few cases, messages in eDialogue developed into a need for phone calls or an extra checkup in the outpatient clinic. The time of the phone call or attendance was then arranged through eDialogue. However, digital communication was perceived as adequate in most cases. There were instances where follow-up questions from HCPs were necessary, yet patients quickly felt understood and equally comprehended the answers they received:

> *Although we have not spoken on the phone, I have received sufficient information and I also feel that I have managed to communicate well.* [Mother of patient 1, DC]

A patient from the ACL group described how eDialogue was used as an extra contact for a him to "fully guard" himself. He was in doubt if the photo sent in eDialogue could show his concerns regarding the surgical site clearly enough, and therefore, he contacted his general practitioner and texted the team in eDialogue at the same time:

> *There was a situation where I had sent a message in the morning, and so, I thought I might as well, while there was still phone time at the GP, call to see if he had an available appointment. Then I came to my GP, and actually got exactly the same answer as I received on the phone (eDialogue) an hour later. So, it wasn't something that was needed as such, but now that I had the opportunity, I thought I might as well do it.* [Patient 8, ACL]

No patients expressed feelings of being misunderstood in their communication with HCPs in eDialogue. They experienced digital communication as being sufficient for their needs; however, they reflected on the risk of misunderstandings when communicating via texts:

> *I think it's a much more optimized way of doing it, because I don't need a physical conversation by phone. I'm fine with texting, but obviously there can*

XSL•FO
RenderX

> *be some misunderstandings or something that can go wrong and then you have to call.* [Mother of patient 15, DC]

The use of photos was mentioned as being very important to support texts. A few patients explained that they lacked the possibility of sending and receiving videos; however, they emphasized that it was not a necessity for their use:

> *If I hadn't been able to send photos, then maybe I would have had to explain something visual by phone, and then I would have had to come in for a checkup, and then I would have wasted a whole day.* [Patient 1, ACL]

> *Video could be nice, but then again, the photos could effectively illustrate how the position of her leg is and show how much she has actually been able to stretch, in what positions it hurts, and so on.* [Mother of patient 17, DC]

The mother of a minor patient explained how she used eDialogue as a photo diary to keep the HCPs across sectors updated on the progress of her son's surgical wound:

> *So, when she (the home care nurse) came and changed the dressings, we took some photos before she put on new ones, and then we kind of had it (photos) from time to time and could follow how it progressed...It was smart as hell, and when it wasn't the same home care nurse coming by, we showed them the photos and at the same time kept the surgeon at the hospital up to date.* [Mother of patient 15, DC]

Use of eDialogue 2 Months After Discharge

The need for support and communication for both patient groups after discharge was expressed through the actual use of eDialogue (Table 5).

Table 5. Patients' and parents' use of eDialogue 2 months after hospital discharge.

	Total number of messages, n	Average number of messages per patient, n	Maximum number of messages per patient, n
DC[a] (n=17)			
All text messages exchanged[b]	338	19.9	54
Text messages sent by patients	189	11.2	34
Actual questions that needed a reply[c]	128	7.5	20
Photos sent by patients[d]	127	7.5	53
ACL[e] (n=14)			
All text messages exchanged	126	9.0	36
Text messages sent by patients	68	4.9	19
Actual questions that needed a reply[f]	55	3.9	14
Photos sent by patients[c]	13	0.9	6

[a]DC: deformity correction.

[b]The total number of text messages exchanged between patients and health care professionals (HCPs) 2 months after discharge.

[c]Text messages sent from the patient or their parents to the HCPs in eDialogue. The minimum number of messages or photos sent per patient was 0, as some patients did not use eDialogue at all.

[d]Actual questions that needed a reply from the HCPs are the number of individual text messages from patients or parents that were formulated as a question; thus, this does not include the back-and-forth 2-way communication that 1 question could lead to (eg, saying thank you).

[e]ACL: anterior cruciate ligament.

[f]Photos refer to the number of photos taken by the patients or parents and sent for review by the HCPs.

Of the patients or their parents, 88% (15/17) in the DC group and 79% (11/14) in the ACL group used eDialogue to ask questions to HCPs after discharge. In the DC group, 13 (87%) of the 15 active users used photos, and in the ACL group, 5 (45%) of the 11 active users sent photos to support communication. Upon inclusion in the study, the patients and the parents were informed that they could expect a response time of 24 hours during the weekdays. This was complied with in 96.2% (176/183) of the cases where a message that required a response from HCPs was sent, and the distribution was equal across groups.

Among users of eDialogue in the DC group, the minimum number of per-patient questions that needed a reply from HCPs was 2, and the maximum was 20. For the ACL group, there was a minimum of 1 and a maximum of 14 questions that needed a reply in 1 dialogue. Thus, there was a marked difference in the individual's use of eDialogue during the study period in both groups.

Most of the communication took place from Monday to Friday; thus, 84.7% (155/183) of the questions that needed a reply from the HCPs were sent and replied to during the weekdays.

The patients and the parents in the DC group used eDialogue throughout the 2 months (Figures 2 and 3), and 15 (88%) of the

XSL•FO
RenderX

17 patients requested to keep on using it after the data collection stopped at 2 months. The patients in the ACL group primarily used eDialogue for the first 2 to 3 weeks after discharge (Figures 2 and 3), and use then faded. Only 2 (6%) of the 31 patients or parents expressed a need to continue with eDialogue after 2 months.

Content analysis of the messages in eDialogue revealed 9 overall categories, including treatment-related issues, rehabilitation and restrictions, concerns about symptoms and complications, medication, psychological support, interdisciplinary and cross-sectoral dialogue, coordination and practical needs, updates and gratitude, HCP ask for feedback. The categories were identified across groups; however, some categories were more prominent in one group than the other (Multimedia Appendix 2).

Figure 2. The number of individual text messages sent from patients or parents to the health care professionals in eDialogue per week 8 weeks after discharge. ACL: anterior cruciate ligament; DC: deformity correction.

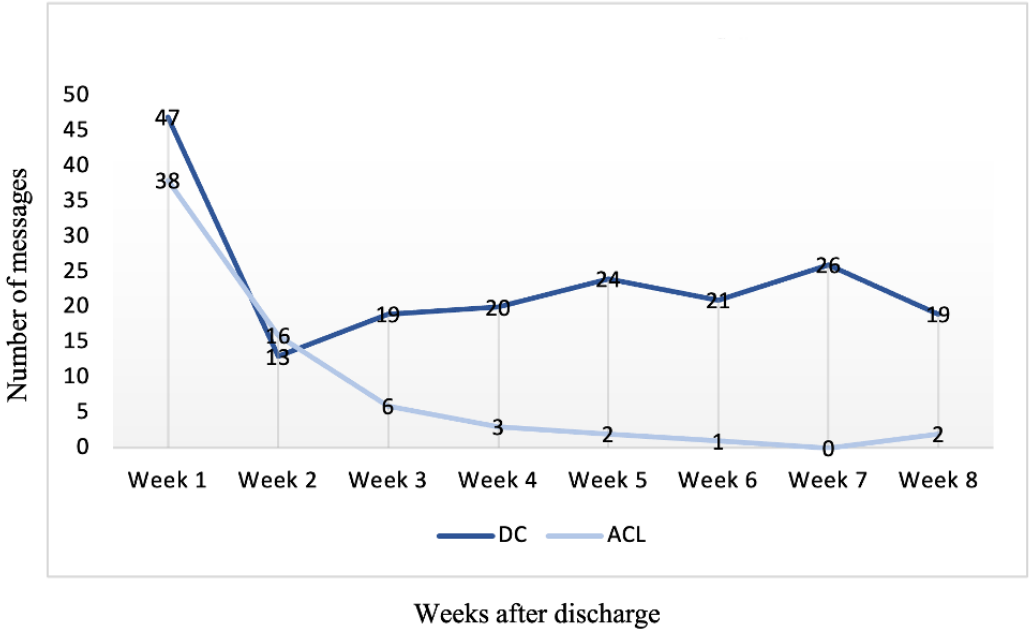
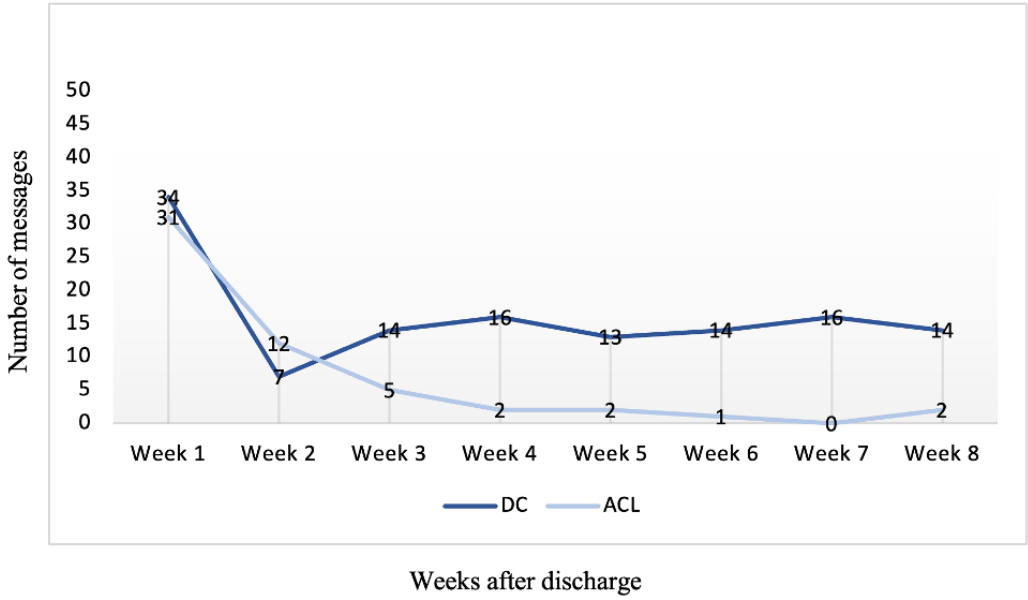


Figure 3. The number of messages sent by patients or parents that required a response from health care professionals, that is, messages phrased as a question, per week 8 weeks after discharge. ACL: anterior cruciate ligament; DC: deformity correction.



XSL•FO
RenderX

Discussion

Principal Findings

Overview

This study explored the perspectives of patients undergoing orthopedic surgery on current communication pathways (phase 1), and their subsequent experiences of using eDialogue after discharge, as well as the actual use of the solution (phase 2).

In phase 1, we identified unmet needs among patients regarding communication with HCPs after discharge. The themes involved perspectives of difficult communication pathways, lack of information due to inadequate coordination of care, and that relation and communication provide "peace of mind." In phase 2, the participants were set up to use eDialogue for 2 months after surgery and discharge, providing them access to direct digital communication with their individual health care team across settings. Through follow-up interviews, they articulated the following themes: digitally enhanced coherence and proximity, drivers and barriers to use, and that eDialogue rearranges communication pathways. Use data of eDialogue supported the experiences expressed in the interviews and provided an overview of the actual use. These findings will be discussed with the theoretical framework of COC and previous research.

Signs of Improved COC With eDialogue

Through initial interviews, the patients and the parents expressed a need for more clear communication pathways after discharge. A patient expressed that it felt like being in a "no man's land." As such, they lacked communicative support at home as well as optimized sharing of knowledge between the HCPs involved in their treatment and care across settings, indicating that informational and management COC is under pressure [32]. Similar findings are described in other studies on patients' experiences of the transition from hospital to home following surgery [24,28], and this emphasizes the need to address communicative challenges around hospital discharge.

The patients and the parents in complex and long-term orthopedic treatments (DC) experienced a greater need for continuous contact with their known health care team than those undergoing day surgery (ACL). Thus, the relationship, trust, and mutual understanding with the HCPs were described as being of great importance for their experience of security. For these patients, access to eDialogue was particularly useful, suggesting that eDialogue may play a role in facilitating relational COC. The patients in the ACL group, despite still having an unmet need for information, expressed that "less" would have been suitable for them. As digital communication becomes more prevalent in health care [1,2,4,5,7,9], comprehensive evaluations are crucial, including efficiency and optimal resource use considerations. Some patients may find less resource-intensive options, such as automated text message interventions, sufficient [36].

Through follow-up interviews, the patients and the parents across groups highlighted that eDialogue provided easy access to relevant HCPs and facilitated coherence and proximity after returning home, leading to "a sense of security." These findings corroborate previous studies [14,37] and support our assumption that team-based digital communication may contribute to improving patients' experiences of COC in transitions from hospital to home [32]. Other studies have also highlighted that COC is one of the factors that can be positively influenced by the use of team-based digital communication [15,16]. Voruganti et al [15] evaluated the feasibility of integrating a web-based communication tool for collaborative care in a pilot randomized controlled trial and found evidence indicating an increase in COC scores in the intervention group; however, the study was unpowered to show the effect statistically. Another study by Lindkvist et al [16] described how access to and use of an eHealth device for text-based communication, image exchange, and data reports between HCPs and parents of preterm infants or pediatric surgery was experienced positively in the transfer period from hospital to home. Moreover, they reported that parents felt it gave a sense of "shared responsibility," which was also expressed by the patients and parents in this study. Thus, they highlighted that eDialogue facilitated the sharing of information, so they no longer had to be the ones passing on information and knowledge between HCPs. This was a role that they often disliked or mistrusted that they could fulfill adequately. The findings from this study indicate, in line with other studies [14-16], that digital team-based communication has the potential to set the framework for interdisciplinary and cross-sector collaboration that supports COC following hospital discharge. Whether team-based digital communication can actually enhance levels of COC to an extent where it can be measured remains to be investigated.

Patients Want to Communicate Digitally

As seen in other studies on digital asynchronous communication [15,16,38], use data demonstrated that most patients and parents across groups used eDialogue (26/31, 84%). The drivers to use eDialogue involved that the tool was recognizable and easy to use. Employing a messenger-like tool, made available to patients on their own smartphone, was a strength, as we did not encounter technical challenges as described in other studies, where devices were newly developed and delivered to participants [16]. The simple solution only allowed for communication in text and photos, and it may lack other options for patients who cannot use the text-based medium. Although previous studies involving text-based digital communication for health care purposes show that patients largely adopt this form of communication across settings and needs [4,10,37,39], digital inclusion in eHealth interventions is important to acknowledge both in regard to the hardware as well as patients' ability to use the solutions [40]. As such, if the patients cannot use the tool, no value has been added. Other studies have integrated several means of communication into their solutions, including text, video, photos, and voice recordings, and found that video communication was especially useful [16,41,42]. This is in contrast to our findings, where patients expressed that the text-based medium was sufficient for them in the postoperative period. However, we acknowledge that eDialogue, as used in this study, may not be sufficient for all patients. When designing and implementing digital communication solutions, considering patients' literacy and eHealth literacy becomes

XSL•FO
RenderX

crucial to ensure equal access to health care [40,43]. Integrating multiple communication modalities within a single solution could serve as a means to achieve this goal.

A driver mentioned in this study was the informality of the solution, and that it felt less interrupting than calling by phone. Similar results have been found in other studies of digital text-based communication [16,37], and this indicates a high degree of acceptance and usability of the solution from the patients' perspectives. With an increasing level of smartphone use in the general population, digital communication becomes a more natural choice when patients need to contact providers. Thus, statistics show that the use of smartphones worldwide is increasing significantly, and in Denmark, it is estimated that 90% of all households own a smartphone [44]. As a barrier to use, the patients expressed concerns about wasting the HCPs' time. This is important to consider when implementing solutions for digital communication. Our findings indicate that this may be offset by a more inviting approach from the HCPs, as some patients and parents expressed this as a facilitator to their use. Previous studies have pointed out the importance of clearly communicating response times when using digital communication [16,37]. Similar findings were highlighted in follow-up interviews of this study, where patients and parents described quick response times, or alternatively late responses, as a driver and a barrier, respectively.

Across groups, the patients and the parents expressed that eDialogue, despite only being an addition to existing communication channels, had rearranged the communication pathways significantly. This became obvious as the patients and parents described a reduced need to call the hospital, as they found eDialogue adequate and exhaustive for their needs. These findings corroborate previous studies showing a potential decrease in phone calls to the hospital after discharge when digital communication is being used [16,45]. By contrast, another study reported, in line with our study, that some questions asked by patients in a digital communication tool were not something they would have called about and thereby indicate that access to digital communication may contribute to an increased consumption of health care resources [16]. To evaluate the effect on resource use, a randomized controlled trial should be performed. Future studies designed to demonstrate the effects on health resource use are desired to shed light on whether digital communication actually reduces patients' use of other forms of communication channels or adds on. In addition, it should be considered whether digital communication provides better quality, for example, defined as COC, patient satisfaction, and security for patients.

This study adds to the knowledge of patients' perspectives on current communication pathways and the sparse evidence of their experiences and use of digital team-based communication, specifically in an orthopedic surgery setting. This may inform future interventions of team-based digital communication, from its application in clinical practice to organizational and management levels.

Limitations

The study has limitations that may affect the interpretation of our results. First, inclusion criteria were participants who owned and used a smartphone and could speak and write Danish well enough to send text messages. Second, we explored the perspectives of 2 selected groups of patients undergoing orthopedic surgery. Therefore, the external validity of the results is unknown for other groups of patients undergoing orthopedic surgery, than the ones we explored.

In planning the study, we decided that initial interviews with patients and parents in phase 1, who were subsequently recruited to use eDialogue after discharge, were appropriate to identify patients' perspectives on current communication pathways. However, some patients found it difficult to express themselves about this, as they had no or little previous experience of an orthopedic surgery context. In addition, there was a risk that the use of initial interviews combined with follow-up interviews within a short timespan (2 months) may have influenced the patients' expressed attitudes in favor of the intervention in the follow-up interviews. Reflecting on this, it might have been better to perform initial interviews with a group of patients who were not given access to eDialogue afterwards.

In this study, we did not use log files to summarize the use data, as other studies have done [16,26], and this may be perceived as a limitation. However, we argue that log files, which report the number of log-in attempts, database entries, messages sent in total and the like, would not show the actual use as it presented to the participants in clinical practice. Therefore, manual counts were used to remove messages saying "thank you" or similar, as these are not considered relevant to the use of eDialogue in a health care setting.

Overall, the 24-hour weekday response time was met in this study and some patients reported extremely fast responses from HCPs. This finding must be interpreted with caution, as we cannot rule out that it is due to the Hawthorne effect, which suggests that people behave better when they are observed [46]. Conversely, it can also be an expression of the flexibility that lies in the digital asynchronous form of communication, giving HCPs the possibility to answer when they have the time for it, or it may simply reflect that the HCPs replied instantly (when able to) not to forget it. Nevertheless, an exclusively positive interpretation of compliance with the response time in this study may result in blindness toward the possible pitfalls that can occur in the real world if eDialogue is implemented. Insights from the perspective of HCPs can reveal this.

Conclusions

The findings from this study indicate that the patients and the parents experienced an unmet need related to communication and collaboration following hospital discharge. eDialogue was overall evaluated positively, and the patients and parents perceived team-based digital communication as correspondent to their needs and suggested that it provided a sense of security after returning home. COC may be enhanced by assembling the team of HCPs in a simple digital communication solution with patients. However, eDialogue should be further evaluated and tested. Future research has to explore HCPs' perspectives on the solution as well as establish the effects and organizational and economic incentives to use team-based digital communication in the context of orthopedic surgery care pathways.

XSL•FO
RenderX

Acknowledgments

The authors would like to thank all patients, parents, health care professionals, and management from departments across hospitals and municipalities in Northern Jutland, Denmark, who agreed to participate in this study. The study was partly funded by the Research Fund of Danish Health Region North (2021-0019). Visma, the company who owns the technical solution used in this study ("LetDialog"), had no role in the design of the study; collection, analyses, or interpretation of data; or in the writing of the manuscript or the decision to publish the results.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Interview guides.
[DOCX File , 16 KB - humanfactors_v11i1e49696_app1.docx]

Multimedia Appendix 2
The number of patients from the deformity correction (DC) and anterior cruciate ligament (ACL) group who sent messages within the respective categories. HCP: health care professional.
[PNG File , 88 KB - humanfactors_v11i1e49696_app2.png]

References

1. Hamza M, Khan HS, Sattar ZA, Hanif M. Doctor-patient communication in surgical practice during the coronavirus (COVID-19) pandemic. Br J Surg 2020 Jun;107(7):e193 [FREE Full text] [doi: 10.1002/bjs.11661] [Medline: 32364265]
2. Cronin RM, Davis SE, Shenson JA, Chen Q, Rosenbloom ST, Jackson GP. Growth of secure messaging through a patient portal as a form of outpatient interaction across clinical specialties. Appl Clin Inform 2015;6(2):288-304 [FREE Full text] [doi: 10.4338/ACI-2014-12-RA-0117] [Medline: 26171076]
3. Using e-health and information technology to improve health. World Health Organization. URL: https://www.who.int/westernpacific/activities/using-e-health-and-information-technology-to-improve-health [accessed 2024-03-21]
4. Newhouse N, Lupiáñez-Villanueva F, Codagnone C, Atherton H. Patient use of email for health care communication purposes across 14 European countries: an analysis of users according to demographic and health-related factors. J Med Internet Res 2015 Mar 06;17(3):e58 [FREE Full text] [doi: 10.2196/jmir.3700] [Medline: 25798912]
5. Baer D. Patient-physician e-mail communication: the kaiser permanente experience. J Oncol Pract 2011 Jul;7(4):230-233 [FREE Full text] [doi: 10.1200/JOP.2011.000323] [Medline: 22043186]
6. McAlearney AS, Sieck CJ, Gaughan A, Fareed N, Volney J, Huerta TR. Patients' perceptions of portal use across care settings: qualitative study. J Med Internet Res 2019 Jun 06;21(6):e13126 [FREE Full text] [doi: 10.2196/13126] [Medline: 31172960]
7. Semple JL, Armstrong KA. Mobile applications for postoperative monitoring after discharge. CMAJ 2017 Jan 09;189(1):E22-E24 [FREE Full text] [doi: 10.1503/cmaj.160195] [Medline: 27920015]
8. Mehta SJ, Hume E, Troxel AB, Reitz C, Norton L, Lacko H, et al. Effect of remote monitoring on discharge to home, return to activity, and rehospitalization after hip and knee arthroplasty: a randomized clinical trial. JAMA Netw Open 2020 Dec 01;3(12):e2028328 [FREE Full text] [doi: 10.1001/jamanetworkopen.2020.28328] [Medline: 33346847]
9. Danbjørg DB, Wagner L, Kristensen BR, Clemensen J. Intervention among new parents followed up by an interview study exploring their experiences of telemedicine after early postnatal discharge. Midwifery 2015 Jun;31(6):574-581. [doi: 10.1016/j.midw.2015.02.007] [Medline: 25765743]
10. Garne Holm K, Brødsgaard A, Zachariassen G, Smith AC, Clemensen J. Parent perspectives of neonatal tele-homecare: a qualitative study. J Telemed Telecare 2019 May;25(4):221-229. [doi: 10.1177/1357633X18765059] [Medline: 29792079]
11. Atherton H, Sawmynaden P, Sheikh A, Majeed A, Car J. Email for clinical communication between patients/caregivers and healthcare professionals. Cochrane Database Syst Rev 2012;11:CD007978. [doi: 10.1002/14651858.CD007978.pub2] [Medline: 23152249]
12. de JC, Ros WJ, Schrijvers G. The effects on health behavior and health outcomes of internet-based asynchronous communication between health providers and patients with a chronic condition: a systematic review. J Med Internet Res 2014;16(1):e19 [FREE Full text] [doi: 10.2196/jmir.3000] [Medline: 24434570]
13. Wibe T, Hellesø R, Varsi C, Ruland C, Ekstedt M. How does an online patient-nurse communication service meet the information needs of men with recently diagnosed testicular cancer? ISRN Nurs 2012;2012:260975 [FREE Full text] [doi: 10.5402/2012/260975] [Medline: 23251816]

XSL•FO
RenderX

14. Ekstedt M, Børøsund E, Svenningsen IK, Ruland CM. Reducing errors through a web-based self-management support system. Stud Health Technol Inform 2014;201:328-334. [Medline: 24943563]
15. Voruganti T, Grunfeld E, Jamieson T, Kurahashi AM, Lokuge B, Krzyzanowska MK, et al. My team of care study: a pilot randomized controlled trial of a web-based communication tool for collaborative care in patients with advanced cancer. J Med Internet Res 2017 Jul 18;19(7):e219 [FREE Full text] [doi: 10.2196/jmir.7421] [Medline: 28720558]
16. Lindkvist RM, Sjöström-Strand A, Landgren K, Johnsson BA, Stenström P, Hallström IK. "In a way we took the hospital home"-a descriptive mixed-methods study of parents' usage and experiences of eHealth for self-management after hospital discharge due to pediatric surgery or preterm birth. Int J Environ Res Public Health 2021 Jun 15;18(12):6480 [FREE Full text] [doi: 10.3390/ijerph18126480] [Medline: 34203985]
17. Continuity and coordination of care: a practice brief to support implementation of the WHO framework on integrated people-centred health services. World Health Organization. 2018. URL: https://www.who.int/publications/i/item/9789241514033 [accessed 2024-03-21]
18. Naylor MD, Aiken LH, Kurtzman ET, Olds DM, Hirschman KB. The care span: the importance of transitional care in achieving health reform. Health Aff (Millwood) 2011 Apr;30(4):746-754 [FREE Full text] [doi: 10.1377/hlthaff.2011.0041] [Medline: 21471497]
19. Forster AJ, Murff HJ, Peterson JF, Gandhi TK, Bates DW. The incidence and severity of adverse events affecting patients after discharge from the hospital. Ann Intern Med 2003 Feb 4;138(3):161-167. [Medline: 12558354]
20. Coleman EA, Parry C, Chalmers S, Min S. The care transitions intervention: results of a randomized controlled trial. Arch Intern Med 2006 Sep 25;166(17):1822-1828. [doi: 10.1001/archinte.166.17.1822] [Medline: 17000937]
21. Goodwin JS. Continuity of care matters in all health care settings. JAMA Netw Open 2021 Mar 01;4(3):e213842 [FREE Full text] [doi: 10.1001/jamanetworkopen.2021.3842] [Medline: 33760086]
22. Oslock WM, Satiani B, Way DP, Tamer RM, Maurer J, Hawley JD, et al. A contemporary reassessment of the US surgical workforce through 2050 predicts continued shortages and increased productivity demands. Am J Surg 2022 Jan;223(1):28-35. [doi: 10.1016/j.amjsurg.2021.07.033] [Medline: 34376275]
23. Wainwright T, Middleton R. An orthopaedic enhanced recovery pathway. Curr Anaesth Crit Care 2010 Jun;21(3):114-120. [doi: 10.1016/j.cacc.2010.01.003]
24. Halding AG, Hovland G, Johansen VA, Hege Eide P, Åslaug Kringeland T. A planned but vulnerable pathway towards self-management following day surgery. Orthopaedic patients' experiences. Scand J Caring Sci 2021 Dec;35(4):1114-1122. [doi: 10.1111/scs.12926] [Medline: 33128304]
25. Jensen LW, Madsen NA, Dinesen B. Developing a digital platform for telerehabilitation of patients treated with external fixation device after complex tibia fractures. In: Proceedings of the 17th Scandinavian Conference on Health Informatics. 2019 Presented at: SHI '19; November 12-13, 2019; Oslo, Norway p. 154-159 URL: https://vbn.aau.dk/en/publications/developing-a-digital-platform-for-telerehabilitation-of-patients-
26. Husain A, Cohen E, Dubrowski R, Jamieson T, Kurahashi AM, Lokuge B, et al. A clinical communication tool (Loop) for team-based care in pediatric and adult care settings: hybrid mixed methods implementation study. J Med Internet Res 2021 Mar 03;23(3):e25505 [FREE Full text] [doi: 10.2196/25505] [Medline: 33656445]
27. Gulmans J, Vollenbroek-Hutten M, van Gemert-Pijnen L, van Harten W. A web-based communication system for integrated care in cerebral palsy: experienced contribution to parent-professional communication. Int J Integr Care 2012;12:e9 [FREE Full text] [doi: 10.5334/ijic.672] [Medline: 22977435]
28. Balakrishnan AS, Palmer NR, Odisho AY. "Can you hear me now?": postoperative patient-initiated communication with providers. Am J Manag Care 2020 Oct 01;26(10):e333-e341 [FREE Full text] [doi: 10.37765/ajmc.2020.88507] [Medline: 33094946]
29. Kom tættere på dine borgere med sikker SMS: first agenda. LetDialog. URL: https://letdialog.dk/ [accessed 2023-02-02]
30. DeWalt KM, DeWalt BR. Participant Observation: A Guide for Fieldworkers. Walnut Creek, CA: AltaMira Press; 2011.
31. Brinkmann S, Kvale S. InterViews: Learning the Craft of Qualitative Research Interviewing. København, Denmark: SAGE Publications; 2015.
32. Haggerty JL, Reid RJ, Freeman GK, Starfield BH, Adair CE, McKendry R. Continuity of care: a multidisciplinary review. BMJ 2003 Nov 22;327(7425):1219-1221 [FREE Full text] [doi: 10.1136/bmj.327.7425.1219] [Medline: 14630762]
33. Bahr SJ, Weiss ME. Clarifying model for continuity of care: a concept analysis. Int J Nurs Pract 2019 Apr;25(2):e12704. [doi: 10.1111/ijn.12704] [Medline: 30393894]
34. Reid R, Haggerty J, McKendry R. Defusing the confusion: concepts and measures of continuity of healthcare. Canadian Health Services Research Foundation. 2002. URL: https://www.researchgate.net/publication/245856177_Defusing_the_Confusion_Concepts_and_Measures_of_Continuity_of_Health_Care#fullTextFileContent [accessed 2024-03-23]
35. Tong A, Sainsbury P, Craig J. Consolidated criteria for reporting qualitative research (COREQ): a 32-item checklist for interviews and focus groups. Int J Qual Health Care 2007 Dec;19(6):349-357 [FREE Full text] [doi: 10.1093/intqhc/mzm042] [Medline: 17872937]

XSL•FO
RenderX

36. Gordon CR, Rezzadeh KS, Li A, Vardanian A, Zelken J, Shores JT, et al. Digital mobile technology facilitates HIPAA-sensitive perioperative messaging, improves physician-patient communication, and streamlines patient care. Patient Saf Surg 2015;9:21 [FREE Full text] [doi: 10.1186/s13037-015-0070-9] [Medline: 26136830]
37. Alexander KE, Ogle T, Hoberg H, Linley L, Bradford N. Patient preferences for using technology in communication about symptoms post hospital discharge. BMC Health Serv Res 2021 Feb 15;21(1):141 [FREE Full text] [doi: 10.1186/s12913-021-06119-7] [Medline: 33588829]
38. Lau C, Churchill RS, Kim J, Matsen FA, Kim Y. Asynchronous web-based patient-centered home telemedicine system. IEEE Trans Biomed Eng 2002 Dec;49(12):1452-1462. [doi: 10.1109/TBME.2002.805456] [Medline: 12542241]
39. Hefner JL, MacEwan SR, Biltz A, Sieck CJ. Patient portal messaging for care coordination: a qualitative study of perspectives of experienced users with chronic conditions. BMC Fam Pract 2019 May 03;20(1):57 [FREE Full text] [doi: 10.1186/s12875-019-0948-1] [Medline: 31053063]
40. Norman CD, Skinner HA. eHealth literacy: essential skills for consumer health in a networked world. J Med Internet Res 2006 Jun;8(2):e9 [FREE Full text] [doi: 10.2196/jmir.8.2.e9] [Medline: 16867972]
41. DeAntonio JH, Kang HS, Cockrell HC, Rothstein W, Oiticica C, Lanning DA. Utilization of a handheld telemedicine device in postoperative pediatric surgical care. J Pediatr Surg 2019 May;54(5):1005-1008. [doi: 10.1016/j.jpedsurg.2019.01.032] [Medline: 30782441]
42. Gund A, Sjöqvist BA, Wigert H, Hentz E, Lindecrantz K, Bry K. A randomized controlled study about the use of eHealth in the home health care of premature infants. BMC Med Inform Decis Mak 2013 Feb 09;13:22 [FREE Full text] [doi: 10.1186/1472-6947-13-22] [Medline: 23394465]
43. WHO guideline: recommendations on digital interventions for health system strengthening. Licence: CC BY-NC-SA 3.0 IGO. World Health Organization. 2019. URL: https://www.who.int/publications/i/item/9789241550505 [accessed 2024-03-21]
44. Tassy A, Berg CE. It-anvendelse i befolkningen 2021. Danmarks Statistik: Forskning, Teknologi og Kultur. 2021. URL: https://www.dst.dk/Site/Dst/Udgivelser/GetPubFile.aspx?id=39431&sid=itbef2021 [accessed 2024-03-21]
45. Armstrong KA, Coyte PC, Brown M, Beber B, Semple JL. Effect of home monitoring via mobile app on the number of in-person visits following ambulatory surgery: a randomized clinical trial. JAMA Surg 2017 Jul 01;152(7):622-627 [FREE Full text] [doi: 10.1001/jamasurg.2017.0111] [Medline: 28329223]
46. Silverman WA. Hawthorne effects. In: Silverman WA, editor. Hawthorne Effects. In Where's The Evidence? Debates in Modern Medicine. New York, NY: Oxford University Press; 1999:58-51.

Abbreviations

ACL: anterior cruciate ligament
COC: continuity of care
DC: deformity correction
HCP: health care professional

Edited by A Kushniruk; submitted 06.06.23; peer-reviewed by N Finn, L Shang; comments to author 24.01.24; revised version received 29.01.24; accepted 20.02.24; published 29.03.24.

Please cite as:
Jensen LWH, Rahbek O, Lauritsen REK, Kold S, Dinesen B
Patient Perspectives on Communication Pathways After Orthopedic Surgery and Discharge and Evaluation of Team-Based Digital Communication: Qualitative Exploratory Study
JMIR Hum Factors 2024;11:e49696
URL: https://humanfactors.jmir.org/2024/1/e49696
doi:10.2196/49696
PMID:38551641

©Lili Worre Høpfner Jensen, Ole Rahbek, Rikke Emilie Kildahl Lauritsen, Søren Kold, Birthe Dinesen. Originally published in JMIR Human Factors (https://humanfactors.jmir.org), 29.03.2024. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on https://humanfactors.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

The Asthma App as a New Way to Promote Responsible Short-Acting Beta2-Agonist Use in People With Asthma: Results of a Mixed Methods Pilot Study

Liselot N van den Berg[1,2], MSc; Cynthia Hallensleben[1,2]; Lisa AE Vlug[1,2], MSc; Niels H Chavannes[1,2], MD, PhD; Anke Versluis[1,2], PhD

[1]Department of Public Health and Primary Care, Leiden University Medical Center, Leiden, Netherlands
[2]National eHealth Living Lab, Leiden, Netherlands

Corresponding Author:
Liselot N van den Berg, MSc
Department of Public Health and Primary Care
Leiden University Medical Center
Hippocratespad 21
Leiden, 2300 RC
Netherlands
Phone: 31 71 526 57 61
Email: L.N.van_den_Berg@lumc.nl

Abstract

Background: Approximately 262 million people worldwide are affected by asthma, and the overuse of reliever medication—specifically, short-acting beta2-agonist (SABA) overuse—is common. This can lead to adverse health effects. A smartphone app, the Asthma app, was developed via a participatory design to help patients gain more insight into their SABA use through monitoring and psychoeducation.

Objective: This pilot study aims to evaluate the feasibility and usability of the app. The preliminary effects of using the app after 3 months on decreasing asthma symptoms and improving quality of life were examined.

Methods: A mixed methods study design was used. Quantitative data were collected using the app. Asthma symptoms (measured using the Control of Allergic Rhinitis and Asthma Test) and the triggers of these symptoms were collected weekly. Quality of life (36-Item Short-Form Health Survey) was assessed at baseline and after 3, 6, and 12 months. User experience (System Usability Scale) was measured at all time points, except for baseline. Furthermore, objective user data were collected, and qualitative interviews, focusing on feasibility and usability, were organized. The interview protocol was based on the Unified Theory of Acceptance and Use of Technology framework. Qualitative data were analyzed using the Framework Method.

Results: The baseline questionnaire was completed by 373 participants. The majority were female (309/373, 82.8%), with a mean age of 46 (SD 15) years, and used, on average, 10 SABA inhalations per week. App usability was rated as good: 82.3 (SD 13.2; N=44) at 3 months. The Control of Allergic Rhinitis and Asthma Test score significantly improved at 3 months (18.5) compared with baseline (14.8; β=.189; SE 0.048; P<.001); however, the obtained score still indicated uncontrolled asthma. At 3 months, there was no significant difference in the quality of life. Owing to the high dropout rate, insufficient data were collected at 6 and 12 months and were, therefore, not further examined. User data showed that 335 users opened the app (250/335, 74.6%, were returning visitors), with an average session time of 1 minute, and SABA registration was most often used (7506/13,081, 57.38%). Qualitative data (from a total of 4 participants; n=2, 50% female) showed that the participants found the app acceptable and clear. Three participants stated that gaining insight into asthma and its triggers was helpful. Two participants no longer used the app because they perceived their asthma as controlled and, therefore, did not use SABA often or only used it regularly based on the advice of the pulmonologist.

Conclusions: The initial findings regarding the app's feasibility and usability are encouraging. However, the notable dropout rate underscores the need for a cautious interpretation of the results. Subsequent studies, particularly those focusing on implementation, should explore the potential integration of the app into standard treatment practices.

(JMIR Hum Factors 2024;11:e54386) doi:10.2196/54386

XSL•FO
RenderX

KEYWORDS

asthma; short-acting beta2-agonist; SABA overuse; app; eHealth; feasibility; usability; mobile phone

Introduction

Asthma is a common chronic inflammatory disease, which is estimated to affect 262 million people worldwide [1]. Step 1 of medical treatment involves the prescription of short-acting beta2-agonist (SABA) as a reliever medication. In contrast to inhaled corticosteroids (ICS), SABA does not have an anti-inflammatory effect on the respiratory tract [2,3]. In 2019, step 1 was modified in the Global Initiative for Asthma guidelines [2]. Specifically, the option of a low dose of ICS-formoterol, as needed, was added because asthma control is often suboptimal [3-5]. According to guidelines, using SABA more than twice a week indicates suboptimal, uncontrolled asthma [3]. Approximately half of the patients with asthma have uncontrolled asthma [5-7]. The overuse of SABA is linked to an increased risk of asthma exacerbations, which are associated with damage to the respiratory tract, asthma-related hospitalization, and visits to the emergency department [8-12].

The overuse of SABA is common for different reasons. First, individuals often overuse their SABA instead of taking ICS to achieve a rapid relief from an asthma attack [13-15]. Second, individuals may lack knowledge about the medication and insight into the actual frequency of medication use [13,16]. For example, the REcognise Asthma and LInk to Symptoms and Experience study [17] found that 80% of the participants thought they had controlled asthma, although 40% had used their SABA ≥3 times during the past week. A post hoc analysis of the Dutch participants from this study showed that 60% of the patients with asthma overused their SABA in the previous week [18].

Previous studies have shown that self-management apps can help reduce the frequency of SABA use, increase SABA-free days, and improve overall asthma control [19,20]. These apps can also boost individuals' confidence in managing asthma and improve their quality of life (QoL) [21-23]. Often, these self-management tools include education, self-monitoring, and feedback to support the end users in managing their disease daily [21,22,24,25]. Most apps are developed using state-based models, such as the Waterfall Model, and agile methods [26]. These traditional methods do not engage end users in the development process, which may result in lower usability and adherence of end users [27]. Therefore, an app was developed in collaboration with end users and other relevant stakeholders (eg, health care professionals) using a participatory design. This design can be used to engage relevant stakeholders during the development process, which may improve the usability and adherence to an app. The objective of the app is to help patients gain more insight into their SABA use while also promoting responsible SABA use. This may eventually decrease SABA overuse. In a previous study, we described the development process of the Asthma app [28].

This pilot study, using a mixed methods design, aims to examine (1) the feasibility and usability of the app in people with asthma and (2) the preliminary effects of using the app after 3 months on decreasing asthma symptoms and improving QoL.

Methods

Design and Population

The pilot study had a mixed methods design. Initially, the study was purely quantitative, with data collected through questionnaires administered in the app to examine the usability and preliminary effects of the app. Individuals were eligible to participate if they (1) were aged ≥18 years and (2) had asthma. Individuals who did not meet these inclusion criteria were excluded from the study; however, they could still use the app. The study period for the participants was 12 months. The study was conducted from January 15, 2021, to December 6, 2022; however, user data were collected until December 31, 2021. User data collection was stopped earlier because the costs for collecting these data increased after 2021, and this could no longer be funded.

During the study, we noticed that most participants used the app only in the first week after downloading. Owing to the high dropout rate, an additional qualitative study was conducted to examine the feasibility and usability of the app in more detail. Individuals who use, had used, or had downloaded the app once were included in the semistructured interviews. Individuals who participated in the qualitative interviews did not necessarily participate in the quantitative study. Qualitative interviews were held until data saturation was reached; data saturation was expected after 6 to 12 interviews [29,30]. Data were collected between November 7, 2022, and December 13, 2022.

Ethical Considerations

According to the Medical Ethics Committee of the Leiden University Medical Center, this study did not fall within the scope of the Dutch Medical Research Involving Human Subjects Act (N20.103). Subsequently, a declaration of no objection was obtained from the Medical Ethics Committee. Participants provided informed consent and were able to opt out (see the *Procedure* section). The quantitative data were collected anonymously, and the qualitative data were collected pseudonymously.

Asthma App

The Asthma app (a Dutch app developed by the Leiden University Medical Center and Innovattic; Figure 1) allows end users to register their SABA use. Moreover, users can register asthma symptoms weekly (they receive a notification to do so), and they can register the triggers of these symptoms at any time. A graph shows how SABA use, asthma symptoms, and their triggers are related. The amount of SABA used was compared with the existing guidelines [2,3] or, when applicable, with health care professional's advice. Psychoeducation is also included, covering topics such as *what is asthma* and *types of medication and their function* [28]. The app was available free of charge in the App Store and Google Play Store.

XSL•FO
RenderX

Figure 1. Visuals of the final version of the Asthma app: (A) landing page where users can register their short-acting beta2-agonist (SABA); (B) the graphical overview or statistics (in Dutch statistieken) where users can get insight into their SABA use, asthma symptoms, and asthma triggers; and (C) psychoeducation or information (in Dutch informatie) where users can learn more about their asthma and the app. On the landing page, users can receive three different messages based on the number of registered SABA compared with the prescription: (1) "You can still use your SABA # times this week," (2) "You are at the maximum recommended dose of SABA for this week," or (3) "You needed more SABA this week than advised." After downloading the app, users receive an explanation on how to interpret the graphical overview, and this explanation can also be found in the informational part of the app.

Procedure

Quantitative Data

Different channels were used to announce the app's go-live and to recruit participants. Relevant organizations (eg, Lung Foundation Netherlands and National eHealth Living Lab) posted the information on their website and social media, or only on their website or social media, and the closed Facebook group *Asthma and Peers* in the Netherlands published the information as well. The information was further communicated through publications (ie, via the COPD Asthma General Practitioners Advice Group in a magazine for pharmacists assistants and in a national newspaper in the Netherlands). Moreover, flyers were distributed via general practices.

After downloading and installing the app, individuals were asked 2 questions to determine their eligibility for the study (ie, whether they were aged ≥18 years and had asthma). Eligible individuals were given information about the study and could decide whether they wanted to participate by signing an informed consent form in the app. In the app, participants could view the informed consent form and withdraw from the study at any moment if they wanted to. If individuals chose to withdraw their consent, they could continue using the app.

Next, participants were asked to complete the demographic and clinical characteristics questionnaire and the baseline questionnaire about QoL (ie, 36-Item Short-Form Health Survey [SF-36] [31]) and intentions to change behavior (ie, a short version of the Theoretical Domains Framework [32]). Asthma symptoms were measured weekly using the Control of Allergic Rhinitis and Asthma Test (CARAT) [33,34]. The triggers of the asthma symptoms, such as dust mites and hay fever, were asked at the end of the CARAT, and the user could also enter additional triggers throughout the week. At 3, 6, and 12 months, user experience (ie, System Usability Scale [SUS]; [35]) and QoL were assessed. No compensation was provided for completing the questionnaire.

Qualitative Data

To gain more insight into the usability and experiences with the app, the following recruitment text was used: "NeLL is looking for (former) users of the Asthma app to get more insight into the usability and experiences with the app, during a one-time interview." We recruited participants for the semistructured interviews via relevant organizations (eg, Asthma Association of the Netherlands and Davos and National eHealth Living Lab) that posted the information on their website and social media, or only on their website or social media; the information was also posted in the closed Facebook group *Asthma and Peers* in

XSL•FO
RenderX

the Netherlands. To increase the interview response rate, participants were recruited via the personal channels of the researchers. When a participant was recruited via personal channels, the researcher did not conduct the interview.

Interested individuals could contact the researchers via email. Subsequently, 1 of the researchers (LNvdB and AEV) would contact them to determine whether they were eligible to participate (ie, aged ≥18 years, having asthma, and [at least] having downloaded the app). Eligible individuals interested in participating received the informed consent form via email. The participants could sign the informed consent form digitally via Castor (ie, a digital, secure research environment) [36]. After signing the informed consent form, the participants received an email invitation to schedule the semistructured interview. We aimed to enroll individuals who use the app and former users (ie, those who had at least downloaded the app).

An interview protocol was developed (Multimedia Appendix 1) based on the Unified Theory of Acceptance and Use of Technology (UTAUT) framework [37]. These interviews were conducted to better understand the perceived usability and feasibility. Interviews were conducted web-based via Microsoft Teams and lasted between 30 and 45 minutes. The participants received a gift card of 30 euros (US $31.2).

Outcome Measures

Demographic and Clinical Characteristics

General information about the participants and their asthma was obtained, including gender, age (birth year), level of education, type of asthma, degree of asthma control, and type of medication. Multiple answers could be selected when answering the question about the type of asthma (ie, allergic asthma, nonallergic asthma, exercise asthma, severe asthma, and do not know) and medication (ie, SABA, ICS, long-acting beta2-agonist (LABA), ICS+LABA, do not know, and no medication use). Furthermore, the participants were asked whether they had received specific advice from their general practitioner on how much SABA they could use per week. When the participant had not received specific advice or did not know whether they had received specific advice, the existing guideline of a maximum of 2 SABA intakes per week was used. When the participants received specific advice from their general practitioner on their SABA use, they could indicate how much SABA they could use per week.

The question “How much SABA did you use last week?” was used as a baseline measure of SABA use. To examine whether an individual’s asthma was stable or unstable during the last week and differed from their average SABA use, an additional question was asked: “How much SABA do you use on average per week?”. In the app, individuals could register their weekly SABA use by clicking on the plus sign shown on the home screen.

The Intention to Change Behavior

The intention to change behavior was assessed using 3 items of the subscale “Intentions” of the Theoretical Domains Framework questionnaire [32]. The original subscale consisted of 4 items, but 1 of the items did not apply to this study and was, therefore, omitted. Items were answered on a 7-point Likert scale ranging from strongly disagree (1) to strongly agree (7). An example of an item is “In the next three months, I intend to use my SABA as prescribed.” A higher score (with a maximum of 21) signified more intent to use their SABA as prescribed in the next 3 months.

Feasibility and Usability

Different types of user data in the app were collected via an analytics platform (ie, PIWIK), namely, (1) which pages are visited in the app (ie, home screen, psychoeducation, user settings page, questionnaires, and the graph) and (2) events (ie, when the app is opened; SABA registrations; number of user clicks on notifications; and, when applicable, made changes in the maximum intake of SABA as advised by the health care professional).

The usability of the app was measured quantitatively using the 10-item SUS [35]. The items were rated on a 5-point Likert scale ranging from strongly disagree (0) to strongly agree (4). The scores were multiplied by 2.5 to obtain the total score ranging from 0 to 100. A higher score indicated that the app was more user-friendly.

A qualitative assessment of the feasibility and usability was conducted through interviews. The interview protocol was based on the UTAUT framework [37], which identified four main factors that influence the intention and use of technology (in this case, an app): (1) performance expectancy, (2) effort expectancy, (3) facilitating conditions, and (4) social influence. Textbox 1 presents an explanation of these factors. Moreover, the UTAUT framework includes four moderating factors: (1) gender, (2) age, (3) experience, and (4) voluntariness of use [37]. These factors and moderating factors were discussed during the interviews.

Textbox 1. Explanation of the factors within the Unified Theory of Acceptance and Use of Technology framework.

Description

- Performance expectancy: the general benefits associated with app use and feasibility of the app
- Effort expectancy: ease of use and usability of the app
- Facilitating conditions: having sufficient resources and knowledge to use the app
- Social influence: the influence of other people (eg, family, friends, and acquaintances) to start and keep using the app and whether they would recommend the app to others

XSL•FO
RenderX

Preliminary Effects

Asthma symptoms and triggers of these symptoms were measured using the 10-item CARAT [33,34]. An example item is "During the last week, because of your asthma/rhinitis/allergy, how many times, on average, did you experience sneezing?." Items were rated on a 4-point scale ranging from 0 (never) to 3 (almost every day). All items were reverse scored, and the total score ranged from 0 (minimal control) to 30 (maximum control). A score of 24 or higher indicated controlled asthma. In addition to the total score, 2 subscales were calculated: a score of the upper airway and a score of the lower airway. The upper airway score ranged from 0 (minimal control) to 12 (maximum control), and the lower airway score ranged from 0 (minimal control) to 18 (maximum control). An additional question was added to identify the symptom triggers: dust mites, animals, smoke, weather, hay fever or pollen, air pollution, smells, and exertion or exercise. Participants were able to select multiple triggers.

Participants' health and health-related QoL were measured using the SF-36 [31]. The SF-36 consists of 2 main categories: physical and mental health [38]. Physical health entailed the physical components and consisted of the following subscales: physical functioning (10 items), role limitations due to physical problems (4 items), bodily pain (2 items), and general health perceptions (5 items). Mental health entailed the mental components and consisted of the following subscales: social functioning (2 items), general mental health (5 items), role limitations due to emotional problems (3 items), and vitality (4 items) [31,39]. All items were recoded into scores ranging from 0 (the poorest level of physical or mental health) to 100 (the best level of physical or mental health) [40], with higher scores indicating better health and higher QoL.

Statistical Analysis

All quantitative data were analyzed using SPSS (version 25.0; IBM Corp) [41]. Descriptive analyses (eg, means, SDs, and percentages) were used to describe the demographic and clinical characteristics of the participants, intention to change behavior, user experience, QoL, and user data (eg, frequency of weekly SABA use). A mixed model was used to determine the change in asthma symptoms over time from the first week of using the app to 3 months after baseline. QoL at 3 months was compared with baseline data using Wilcoxon signed rank tests. The effects at 6 and 12 months were not examined because of the high dropout rate during the study period (88.2% at 3 months and more dropouts beyond that).

Interviews were audiotaped for subsequent analyses, and all audio records were transcribed intelligent verbatim by 1 researcher (AEV). Qualitative data analyses were performed by 2 researchers (LNvdB and AEV) according to the principles of the Framework Method [42] using Atlas.ti (version 22.0) [43]. The Framework Method is a systematic and flexible approach often used for the thematic analysis of semistructured interview data. Following transcription, the 2 researchers immersed themselves in the interviews to gain a comprehensive understanding. Subsequently, a deductive approach was adopted to code the interviews based on a predefined concept codebook developed beforehand based on the UTAUT framework [37]. The coding process was conducted independently by the 2 researchers, followed by a comparison of the codes. Additional codes were incorporated into the codebook, where applicable. A framework matrix was used to organize the data comprehensively, featuring relevant quotes from the participants. Finally, the characteristics and distinctions within the data set were identified. Throughout the process, the steps and data were discussed with the researchers CH and AV.

Results

Demographic and Clinical Characteristics

In the quantitative study, 485 individuals participated at baseline. Of these 485 individuals, 373 (76.9%) reported that they used SABA. Only these individuals were included in the analysis. Most of the participants were female (309/373, 82.8%) with a mean age of 46 (SD 15) years, had a secondary vocational education or higher (316/373, 84.7%), and had allergic asthma (187/373, 50.1%). At baseline, participants stated that they used, on average, 10 SABA per week and 10 SABA in the week before using the app. Moreover, the mean intention to change behavior was 17.1. This indicates that the participants wanted to use their SABA as prescribed for the next 3 months. Table 1 shows an overview of the demographic and clinical characteristics and the intention to change behavior.

XSL•FO
RenderX

Table 1. Baseline demographic and clinical characteristics of the participants and the intention to change behavior in the quantitative study.

Characteristic	Values
Gender (n=373), n (%)	
Male	63 (16.9)
Female	309 (82.8)
Rather not say	1 (0.3)
Age[a] (y; n=371), mean (SD; range)	46.1 (15; 18-81)
Educational level (n=373), n (%)	
Primary school	7 (1.9)
Secondary education	50 (13.4)
Secondary vocational education	136 (36.5)
Higher professional education	121 (32.4)
University education	59 (15.8)
Type of asthma[b] (n=373), n (%)	
Allergic asthma	187 (50.1)
Nonallergic asthma	126 (33.8)
Exercise asthma	164 (44)
Severe asthma	104 (27.9)
Do not know	31 (8.3)
Self-reported asthma control (n=373), n (%)	
Good control	134 (35.9)
Insufficient control	151 (40.5)
Do not know	86 (23.1)
Medication type used[b] (n=373), n (%)	
SABA[c,d]	373 (100)
ICS[e]	198 (53.1)
LABA[f]	150 (40.2)
ICS+LABA	127 (34)
Do not know	0 (0)
No medication use	0 (0)
Average SABA use in the last week: self-reported (n=373), mean (SD; range)	10.5 (12.6; 0-60)
Average SABA use per week: self-reported (n=373), mean (SD; range)	9.7 (11.6; 0-60)
Had medication advice from the health care professional (n=373), n (%)	
Yes	246 (66)
No	110 (29.5)
Do not know	17 (4.6)
Average maximum prescribed SABA[g] (n=246), mean (SD; range)	22.2 (16.9; 0-60)
Intention to change behavior (n=373), mean (SD; range)	17.1 (4.5, 3-21)

[a]The birth year of 2 participants was missing. These participants were excluded from the calculation of the mean age.
[b]Participants were able to select multiple answers.
[c]SABA: short-acting beta2-agonist.
[d]51 participants only used SABA and no other inhalers.
[e]ICS: inhaled corticosteroids.

XSL•FO
RenderX

[f]LABA: long-acting beta2-agonist.

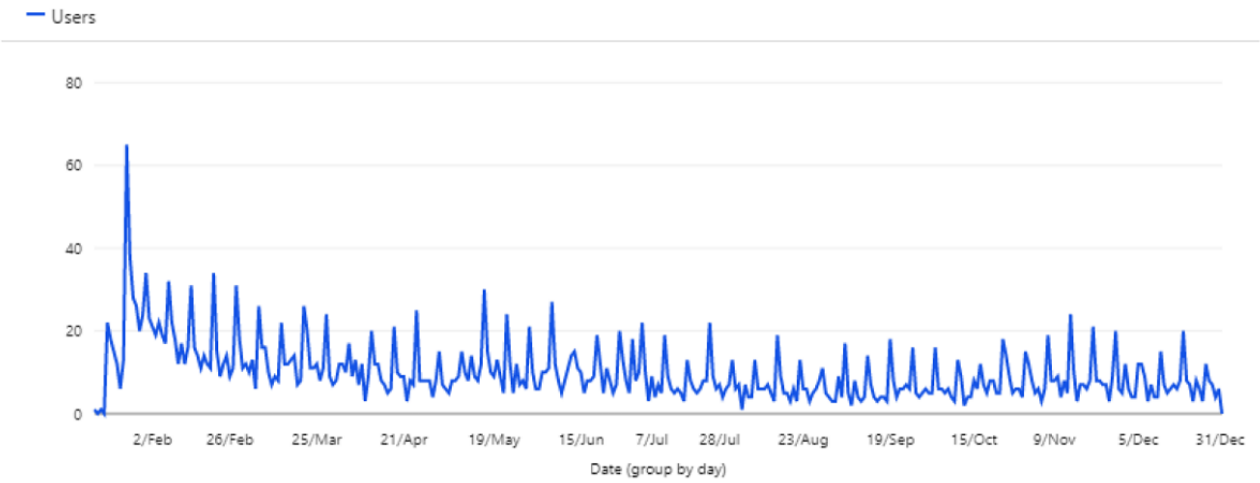
[g]The maximum number of SABA inhalations per week, as prescribed by the participant's health care professional.

In the qualitative part of the study, among the 6 to 12 participants that we planned to recruit, only 4 participants could be included and interviewed. Half of the interviewed participants were female (2/4, 50%), with a mean age of 55 (range 21-78) years. One participant completed senior general secondary education, 1 completed secondary vocational education, and 2 had higher professional education. Two participants stated that they still used the app: they had both been using it for 1 year and 5 months. Regarding *social influence* from the UTAUT framework [37], the app was recommended by the hospital to one participant, and the other participant found it via the asthma association. Two participants stated that they no longer used the app but had used it for approximately 1 or 2 weeks. They both started using the app after the recommendation from a family member.

Feasibility and Usability

User data showed that 335 unique users opened the app, of which 250 (74.6%) were returning visitors, with an average session time of 1 minute. An overview of the number of users during the study period is shown in Figure 2.

Figure 2. Number of users during the study period (January 15, 2021, to December 31, 2021).



Most users opened the app via their smartphone (303/335, 90.4%), followed by a tablet (27/335, 8.1%) and a phablet (4/335, 1.2%). On average, the users had 5 events (ie, starting the app, adding SABA, removing SABA, changing the maximum amount of SABA, and clicking on 1 of the notifications) per session. Registration of SABA (ie, add-function) was most often used (7506/13,081 times, 57.38%). An overview of the events used per week is shown in Figures 3 and 4. At 3, 6, and 12 months, users registered an average of 5 SABA intakes per week (Table 2).

XSL•FO
RenderX

Figure 3. Unique events used per week: "add" means registering short-acting beta2-agonist (SABA), "remove" means removing a SABA registration, "set-max" means changing the maximum amount of SABA, and "start" means starting the app after giving informed consent and filling in the first questionnaires.

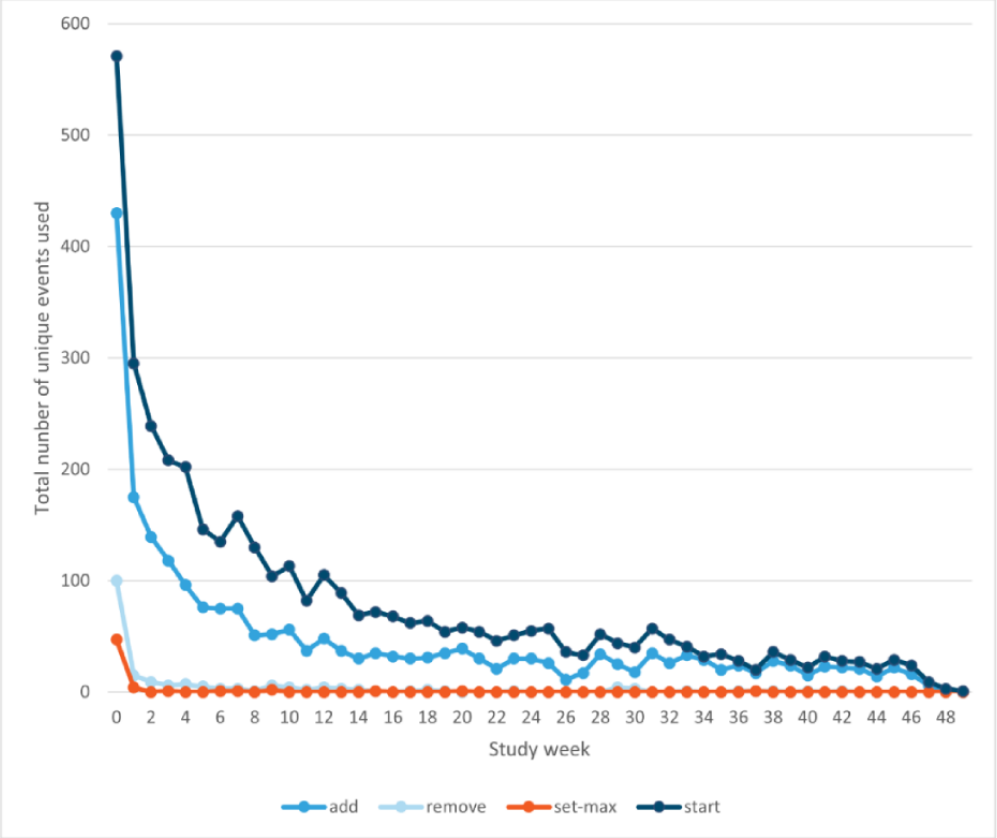
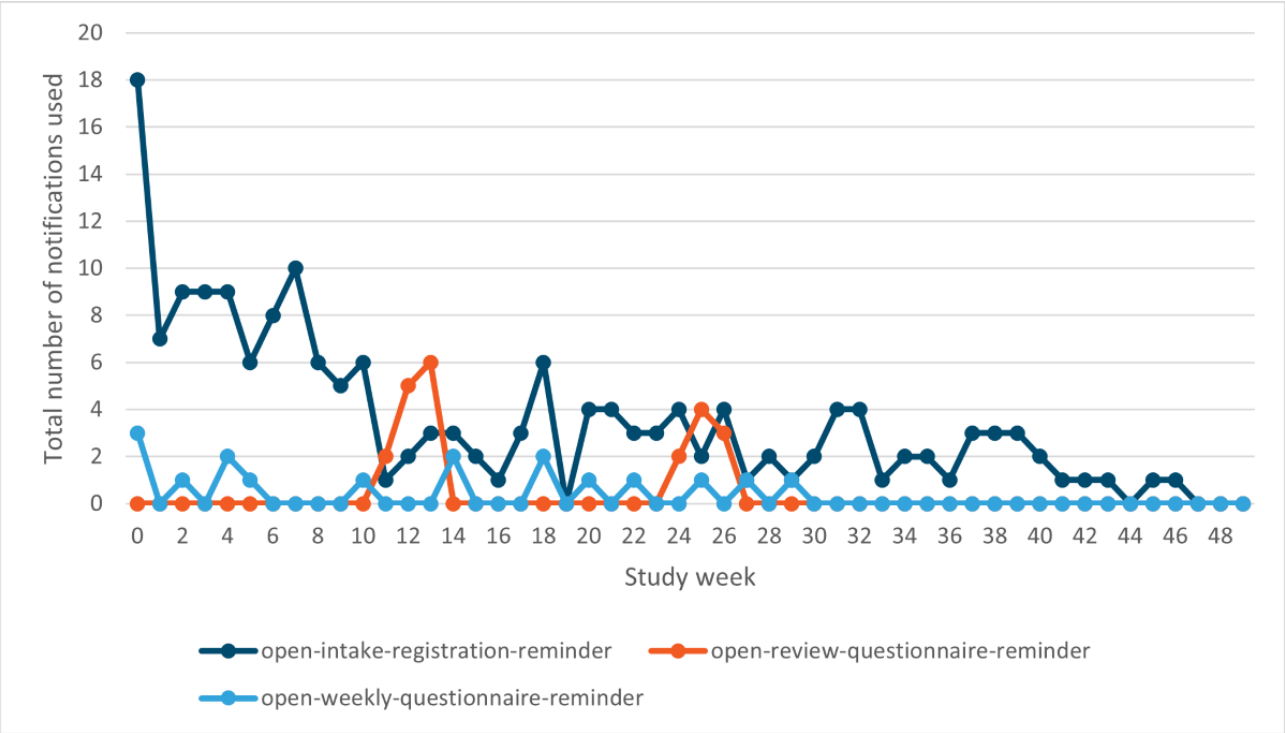


Figure 4. Events used per week, whether participants opened the app via 1 of the notifications: "open-intake-registration-reminder" is the notification users received when they did not register any short-acting beta2-agonist before the end of the week (Sunday); "open-review-questionnaire-reminder" is the notification for the questionnaires used at 3, 6 and 12 months; and "open-weekly questionnaire-reminder" is the notification for the weekly Control of Allergic Rhinitis and Asthma Test.



XSL•FO
RenderX

Table 2. Registered short-acting beta2-agonist use per week.

Time point	Values, mean (SD; 95% CI)
Baseline (n=373)	
Baseline questionnaire	10.5 (0.5; N/A[a])
3 months (n=19)	
12 weeks after baseline	4.67 (2.5; 1.64-13.29)
6 months (n=11)	
25 weeks after baseline	4.82 (2.6; 1.68-13.81)
Latest time point[b] (n=2)	
48 weeks after baseline	5.24 (3; 1.71-16.09)

[a]N/A: not applicable.

[b]Measures ended within 1 year (mid-January 2021 until the end of December 2021).

Usability of the app, as assessed with the SUS, was good over the entire study period: 82.3 (SD 13.2; n=44) at 3 months, 84 (SD 13.6; n=26) at 6 months, and 82.3 (SD 13.4; n=11) at 12 months (Table 3).

Table 3. Questionnaire results regarding usability and quality of life.

Questionnaire	Baseline	3 months	6 months	12 months
Usability				
Values, n (%)	N/A[a]	44 (100)	26 (100)	11 (100)
Values, mean (SD)	N/A	82.3 (13.2)	84 (13.6)	82.3 (13.4)
Quality of life: physical health				
Values, n (%)	373 (100)	44 (100)	26 (100)	11 (100)
Values, mean (SD)	53.6 (22.4)	56.1 (23.9)	57.6 (21.8)	54.9 (24.7)
Quality of life: mental health				
Values, n (%)	373 (100)	44 (100)	26 (100)	11 (100)
Values, mean (SD)	57.4 (21.2)	62.9 (22.3)	59.4 (20.1)	67.8 (18.1)

[a]N/A: not applicable.

Qualitative data showed that 3 (75%) of the 4 participants had experience using other health apps. The users mentioned that they wanted to use health apps that were fun and useful:

> *[...] I only want apps which I like or which are useful.* [Male user, 78 years old]

The participants found the app acceptable and clear in terms of *performance expectancy*. Three participants stated that gaining insight into asthma and its triggers was helpful. Another participant explained that it was not helpful at the moment because he considered his asthma to be controlled. This was for both participants who no longer used the app. One participant did not use SABA often, and the other participant only used it regularly based on the advice of his pulmonologist:

> *First impression was, well, I think, it looks clear. It was pretty clear to me on my own what I could do with it. After using it, yeah, I think it just looks like a nice app, not too old-fashioned. But just fairly new, as you expect from an app in this day and age. And it was also very quickly clear to me exactly what I could do with it.* [Male user, 21 years old]

> *As very useful; you open the app and click on the plus icon how many times if you use it at that time. And also very nice that you get a notification every now and then like, "hey, it is the end of the week; make sure you fill in the amount." Especially if you forget to fill it in. That is nice.* [Male user, 21 years old]

Regarding *effort expectancy* and *facilitating conditions*, 3 participants stated that the app is easy to use and straightforward and does not require much effort to register SABA use. One of these participants also stated that the app was well written and easy to read. The fourth participant did not say anything about ease of use. However, 1 participant experienced difficulties in interpreting the questions and answering the possibilities of the CARAT:

> *So with a few questions, I got, well you already noticed that I have some difficulties with choosing the right one.* [Male user, 78 years old]

Of the participants who continue to use the app (2/4, 50%), they use it multiple times per week, with a minimum frequency of once per week and often 2 or 3 times per week. Opening the app was, for 1 participant, mostly completed after receiving a

XSL•FO
RenderX

notification. A former user mentioned that he would use the app once a week to fill in all the SABA intakes for that week.

Multiple possibilities for improvement were mentioned during the interviews. One participant wanted to be able to fill in triggers that were not listed in the app and also wanted to have the possibility to add more types of medication. Another participant missed contact with other patients with asthma in the app to discuss, for example, medication use. Someone else would change the CARAT based on their experienced difficulties. The last participant missed more background information about SABA use and why SABA should not be used more than twice a week:

> *There is one question that I do not understand. I filled it in good conscience in, and it immediately gave a number that should be decisive, but that I think "yes, but this does not apply to me." In the app, it asked "how many times a night do you wake up?" [...] I do wake up but with a different cause [...] I personally think, but that is my opinion, there should stand "Do you wake up at night, because of your asthma?* [Male user, 78 years old]

Further exploring *social influence* showed that all the participants would recommend the app to others because they experienced that it provided more insight into their medication use and they received more information about the complete picture of asthma. One participant had already recommended the app to an acquaintance, who also started using the app. The 2 former users would specifically recommend it to certain patients: people with severe asthma or uncontrolled asthma or people who do not take their reliever medication as intended. Furthermore, 3 participants also found it useful to show the app to their health care professional during a consultation:

> *I would recommend it, especially to people who do not really have a case like mine. I would also not recommend it to people who, like me, only use salbutamol for sports. Yes, I do not know those people who just do it for sports just like me. Then it does not make much sense to keep track of how often you use it. You just know how often you exercise, and if you first use salbutamol then you know "hey, I use it so often." But for people who use it often, it seems to me that is a very handy app, especially if you can see in that graph how often you have used it per week and in which week more and in which week less.* [Male user, 21 years old]

Preliminary Effects

At week 1, the mean CARAT score was 14.8. This indicated that the participants' asthma was uncontrolled. Their CARAT score improved significantly to a mean score of 18.5 after 12 weeks (ie, 3 months; β=.189; SE 0.048; P<.001); however, this mean score still indicated that their asthma was uncontrolled. This was also the case for both the upper airway score, which significantly improved from a mean score of 6.8 to 7.7 after 12 weeks (β=.073; SE 0.027; *P*=.009), and the lower airway score, which significantly improved from a mean score of 8 to 10.8 after 12 weeks (β=.121; SE 0.037; *P*=.002).

The top three asthma triggers reported in week 1 were (1) weather (321/435, 73.8%), (2) exertion or exercise (305/435, 70.1%), and (3) smoke (197/435, 45.3%). After 12 weeks (ie, 3 months), the top three triggers were (1) weather (25/37, 68%), (2) exertion or exercise (19/37, 51%), and (3) hay fever or pollen (17/37, 46%).

At 3 months, there was no significant difference compared with baseline regarding the mean physical and mental health scores (*Z*=−0.074; *P*=.94 and *Z*=−0.117; *P*=.91, respectively; Table 3).

Discussion

Principal Findings

This study aimed to determine the feasibility and usability of a newly developed app, the Asthma app. Furthermore, the preliminary effects of using the app after 3 months on decreasing asthma symptoms and improving QoL were examined. The quantitative data showed that the usability was good. This was also found in the qualitative data: the app was considered easy to use, and it did not take much effort to register SABA. Furthermore, most participants stated that the app was useful for gaining insight into asthma, triggers, and medication use, and therefore, the app was considered feasible and usable. Multiple improvement possibilities were mentioned during the interviews, such as adding additional personal triggers next to the existing standard list of triggers and the availability of a social support network to contact others with asthma easily. In addition, former users, who no longer used the app, stated that they would recommend the app to people with severe asthma or uncontrolled asthma or people who do not take their reliever medication as intended.

As for the preliminary effects, an improvement in asthma symptoms was found after 3 months; however, the mean asthma symptom score still indicated that the asthma was uncontrolled. Improvement in asthma symptoms was also found in other eHealth studies [19,20]. The mean asthma symptom score in our study, indicating uncontrolled asthma, could be explained by the low intensity and noninvasive nature of the intervention (eg, users could use the app whenever and how often they wanted). A systematic review [44] also found that asthma control did not significantly improve in other studies. They proposed additional well-designed studies to gather more robust findings on what is necessary to achieve optimal asthma control [44]. In terms of QoL, no significant improvement was observed after 3 months. No effect was observed because poor asthma control was associated with worsened QoL [45,46]. The average uncontrolled asthma scores at week 1 and 3 months after baseline can be related to the low QoL scores at the same time points. Moreover, a systematic review [47] demonstrated that eHealth interventions have an inconsistent impact on QoL in people with asthma. The systematic enhancement of clinical outcomes such as QoL was mostly observed within the whole-systems approach, taking into account patient, professional, and organizational elements.

The data from this study should be interpreted with caution because of the high dropout rate, which resulted in insufficient

XSL•FO
RenderX

data for conducting analyses at 6 and 12 months. Although a high dropout rate is frequently seen in studies investigating digital applications, we envisioned that the dropout rate would be lower in this study, considering the participatory design process [28]. The dropout may be explained by the higher probability of dropout in people with chronic diseases when they are impacted physically and mentally by the condition [48]. Most of the participants in this study had uncontrolled asthma and, therefore, more symptoms throughout the day and night. This could have resulted in lower or no app use, which was directly linked to the withdrawal from the study. Another explanation for the high dropout rate could be, as described in our previous study [28], that only a minimal viable product was evaluated. Not all features recommended by the patients, such as registering additional controller medication, were implemented. Therefore, the app might not fit the needs of all the users and cause them to stop using the app.

Using the UTAUT framework [37], performance expectancy was positively associated with the use of the app for the current users. The app will help them gain more insight into asthma, triggers, and medication use. Performance expectancy was lower for former users who stated that their asthma was controlled; therefore, the aim of the app did not align with their needs. Effort expectancy was positively associated with both the intention to use and the actual use of the app, largely because of its user-friendly interface, minimal effort required for SABA registration, and language simplicity. The only aspect that was negatively related to the effort expectancy factor was difficulty with one of the questionnaires by a former user. Facilitating conditions were positively associated with the use of the app. The participants had the appropriate knowledge and resources to use the app. Technical support was not discussed during the interviews; however, clarity regarding the appropriate contact for technical issues could enhance user experience. Finally, social influence played an essential role in intention and use; all interviewees initiated app use through social media discovery or recommendations from health care professionals or family members. They would also recommend the app to others, and 1 participant had already recommended the app to an acquaintance. However, in future studies, this could be further explored in relation to voluntariness of use, which was not thoroughly explored in this study. This is also the case for other moderating factors such as *gender* and *age*. The sample size was too small to explore the associations between the moderating factors, factors, and intention and use of the app. Notably, prior experience with health apps positively influenced the intention and use of the app in this study, and current experience was positively influenced by effort expectancy, facilitating conditions, and social influence for current users.

Strengths and Limitations

This study has several strengths and limitations. A notable strength was the use of a real-life setting for evaluating the Asthma app, allowing a comprehensive understanding of its feasibility and usability. In addition, interviews with both current and former users provided a nuanced perspective on user satisfaction and the factors influencing app use.

In addition to the previously mentioned high dropout rate, another limitation was that the questionnaires were exclusively offered in the app environment. Therefore, former users were no longer able to complete the study questionnaires, thus limiting the availability of their data at later time points (ie, after baseline). To obtain the perspectives of former users on feasibility and usability, they were included in the qualitative interviews. Nevertheless, the recruitment of this group was difficult, and only 2 former users could be included.

Finally, the intended target of 6 to 12 interviews to achieve data saturation [29,30] was not attained. This was partially attributed to the difficulty in reaching former users who may have lost interest in the app or study. Despite the small number of interviews conducted, similar findings were found during data collection between the 2 users and the 2 former users.

Implications for Future Research and Practice

A minimal viable product was examined in this study. During the next development round, feedback gathered during the cocreation of the app could be re-evaluated [28], or new cocreation sessions could be organized to further enhance the app. In future studies, with a newer version of the app, the outcomes of this study could be further examined with more data at more time points, and clinical outcomes, such as the impact of the app on medication adherence, could be explored. A smart asthma inhaler [49,50] could also be linked to the app to gather real-time objective data instead of self-reported registration, which is more sensitive to biases.

This study has a high dropout rate. Renzi et al [51] stated in their review that reminders are often used to improve medication adherence in eHealth interventions but that this improvement is reduced over time. Typically, after 6 months, users tend to revert to their previous behaviors as the novelty of the eHealth intervention wanes [51]. This could also be the case in this study, especially because of the anonymous nature and the use of in-app questionnaires. In future studies, it would be advisable to collect data pseudonymously and send questionnaires via email to achieve a higher response rate. In this way, participants will also be less likely to withdraw from the study and stay involved for longer.

In the new version of the app, additional information about the treatment guidelines should be implemented, such as the fact that users should follow the advice from their health care professional if they receive any. It should be clarified that the app is specifically for people with asthma who only use SABA (and not ICS), which has been the first step of treatment for decades. Potential users could be reached via general practitioners, specialized practice nurses, or pharmacists when they prescribe or distribute SABA. Currently, the Asthma app is a stand-alone app, which means that it is used by patients without the involvement of health care professionals. However, involving health care professionals via “blended care” could improve the quality of care [52]. Moreover, health care professionals can offer additional education and guidance based on the data from the app [53]. To incorporate the app into standard treatment, it is necessary to develop a plan together with asthma associations and health care professionals. A designated implementation team can improve the success rate

XSL•FO
RenderX

of the implementation [54], and it is important to explore context-specific strategies that align with the implementation process phase [55]. Certain barriers (eg, technical issues, time and attention requirements for use, low engagement from health care professionals, and shortage of funding) and facilitators (eg, stakeholder engagement and enthusiasm, minimizing workflow interruptions, and access to information about the app) should be taken into account when implementing the app in standard care [27,56-58]. In addition, more education about SABA overuse could make health care professionals more aware of the risks, which could prioritize the use of the app.

Conclusions

This study evaluated the feasibility and usability of a new app for people with asthma. The initial results regarding usability were positive. Nevertheless, it is essential to exercise caution when interpreting these results because of the high dropout rate in this study. Two former users would recommend the app to people with severe asthma or uncontrolled asthma or people who do not use their reliever medication as intended. Future (implementation) studies could evaluate the potential of incorporating the app into standard treatment practices. Moreover, the actual impact of the app on clinical outcomes, such as medication adherence, should be further examined.

Acknowledgments

This study was funded by Astra Zeneca (no grant number). The authors would like to thank Andrea Sellink and Eva van den Berg who contributed to the review of the manuscript.

Authors' Contributions

LNvdB, CH, NHC, and AV contributed to the conceptualization and reviewed and edited the manuscript. LNvdB and AEV prepared the original draft and conducted the quantitative and qualitative analyses. LNvdB provided the visualization (preparation, creation, and presentation of the published work, specifically visualization or data presentation). LNvdB, CH, and AV contributed to the methodology, validation, investigation, resources, and data curation. CH, NHC, and AV supervised the project and acquired funding for the project. AV was responsible for the project administration. All authors have read and agreed to the final version of the manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Semistructured interview protocol about the Asthma app.
[DOCX File , 15 KB - humanfactors_v11i1e54386_app1.docx]

References

1. The Global Asthma Report 2022. Int J Tuberc Lung Dis 2022 Nov 25;26(Supp 1):1-104. [doi: 10.5588/ijtld.22.1010] [Medline: 36303302]
2. Global strategy for asthma management and prevention. Global Initiative for Asthma. URL: https://ginasthma.org/reports/ [accessed 2020-09-08]
3. Bottema JW, Bouma M, Broekhuizen L, Chavannes NH, Frankemölle LA, Hallensleben C, et al. Astma bij volwassenen (derde herziening). Nederlands Huisartsen Genootschap. 2020 Jul. URL: https://www.nhg.org/standaarden/volledig/nhg-standaard-astma-bij-volwassenen [accessed 2020-08-12]
4. Belhassen M, Nibber A, Van Ganse E, Ryan D, Langlois C, Appiagyei F, et al. Inappropriate asthma therapy-a tale of two countries: a parallel population-based cohort study. NPJ Prim Care Respir Med 2016 Oct 13;26(1):16076 [FREE Full text] [doi: 10.1038/npjpcrm.2016.76] [Medline: 27735927]
5. Demoly P, Annunziata K, Gubba E, Adamek L. Repeated cross-sectional survey of patient-reported asthma control in Europe in the past 5 years. Eur Respir Rev 2012 Mar 01;21(123):66-74 [FREE Full text] [doi: 10.1183/09059180.00008111] [Medline: 22379176]
6. Kosse RC, Murray E, Bouvy ML, de Vries TW, Stevenson F, Koster ES. Potential normalization of an asthma mHealth intervention in community pharmacies: applying a theory-based framework. Res Social Adm Pharm 2020 Feb;16(2):195-201 [FREE Full text] [doi: 10.1016/j.sapharm.2019.05.004] [Medline: 31109820]
7. Kritikos V, Price D, Papi A, Infantino A, Ställberg B, Ryan D, et al. A multinational observational study identifying primary care patients at risk of overestimation of asthma control. NPJ Prim Care Respir Med 2019 Dec 05;29(1):43 [FREE Full text] [doi: 10.1038/s41533-019-0156-4] [Medline: 31804501]
8. Stanford RH, Shah MB, D'Souza AO, Dhamane AD, Schatz M. Short-acting beta-agonist use and its ability to predict future asthma-related outcomes. Ann Allergy Asthma Immunol 2012 Dec;109(6):403-407. [doi: 10.1016/j.anai.2012.08.014] [Medline: 23176877]

XSL•FO
RenderX

9. Nwaru BI, Ekström M, Hasvold P, Wiklund F, Telg G, Janson C. Overuse of short-acting β2-agonists in asthma is associated with increased risk of exacerbation and mortality: a nationwide cohort study of the global SABINA programme. Eur Respir J 2020 Apr 16;55(4):1901872 [FREE Full text] [doi: 10.1183/13993003.01872-2019] [Medline: 31949111]
10. Rennard SI, Farmer SG. Exacerbations and progression of disease in asthma and chronic obstructive pulmonary disease. Proc Am Thorac Soc 2004;1(2):88-92. [doi: 10.1513/pats.2306026] [Medline: 16113418]
11. FitzGerald JM, Tavakoli H, Lynd LD, Al Efraij K, Sadatsafavi M. The impact of inappropriate use of short acting beta agonists in asthma. Respir Med 2017 Oct;131:135-140 [FREE Full text] [doi: 10.1016/j.rmed.2017.08.014] [Medline: 28947020]
12. Anis AH, Lynd LD, Wang XH, King G, Spinelli JJ, Fitzgerald M, et al. Double trouble: impact of inappropriate use of asthma medication on the use of health care resources. CMAJ 2001 Mar 06;164(5):625-631 [FREE Full text] [Medline: 11258208]
13. O'Byrne PM, Jenkins C, Bateman ED. The paradoxes of asthma management: time for a new approach? Eur Respir J 2017 Sep 09;50(3):1701103 [FREE Full text] [doi: 10.1183/13993003.01103-2017] [Medline: 28889114]
14. Partridge MR, van der Molen T, Myrseth SE, Busse WW. Attitudes and actions of asthma patients on regular maintenance therapy: the INSPIRE study. BMC Pulm Med 2006 Jun 13;6(1):13 [FREE Full text] [doi: 10.1186/1471-2466-6-13] [Medline: 16772035]
15. Janson C, Menzies-Gow A, Nan C, Nuevo J, Papi A, Quint JK, et al. SABINA: an overview of short-acting β2-agonist use in asthma in European countries. Adv Ther 2020 Mar 24;37(3):1124-1135 [FREE Full text] [doi: 10.1007/s12325-020-01233-0] [Medline: 31981105]
16. Amin S, Soliman M, McIvor A, Cave A, Cabrera C. Understanding patient perspectives on medication adherence in asthma: a targeted review of qualitative studies. Patient Prefer Adherence 2020 Mar;14:541-551 [FREE Full text] [doi: 10.2147/PPA.S234651] [Medline: 32210541]
17. Price D, Fletcher M, van der Molen T. Asthma control and management in 8,000 European patients: the REcognise Asthma and LInk to Symptoms and Experience (REALISE) survey. NPJ Prim Care Respir Med 2014 Jun 12;24(1):14009 [FREE Full text] [doi: 10.1038/npjpcrm.2014.9] [Medline: 24921985]
18. Baron AJ, Flokstra-de Blok BM, Kerstjens HA, Koopmans-Klein G, Price DB, Sellink AA, et al. High use of SABAs is associated with higher exacerbation rate in Dutch patients with asthma. J Asthma Allergy 2021 Jul;14:851-861 [FREE Full text] [doi: 10.2147/JAA.S292943] [Medline: 34285512]
19. Barrett MA, Humblet O, Marcus JE, Henderson K, Smith T, Eid N, et al. Effect of a mobile health, sensor-driven asthma management platform on asthma control. Ann Allergy Asthma Immunol 2017 Nov;119(5):415-21.e1. [doi: 10.1016/j.anai.2017.08.002] [Medline: 29150069]
20. Merchant RK, Inamdar R, Quade RC. Effectiveness of population health management using the propeller health asthma platform: a randomized clinical trial. J Allergy Clin Immunol Pract 2016;4(3):455-463. [doi: 10.1016/j.jaip.2015.11.022] [Medline: 26778246]
21. Gibson PG, Powell H, Coughlan J, Wilson AJ, Abramson M, Haywood P, et al. Self-management education and regular practitioner review for adults with asthma. Cochrane Database Syst Rev 2003(1):CD001117. [doi: 10.1002/14651858.CD001117] [Medline: 12535399]
22. Pinnock H, Parke HL, Panagioti M, Daines L, Pearce G, Epiphaniou E, et al. Systematic meta-review of supported self-management for asthma: a healthcare perspective. BMC Med 2017 Mar 17;15(1):64 [FREE Full text] [doi: 10.1186/s12916-017-0823-7] [Medline: 28302126]
23. Morrison D, Wyke S, Agur K, Cameron EJ, Docking RI, Mackenzie AM, et al. Digital asthma self-management interventions: a systematic review. J Med Internet Res 2014 Feb 18;16(2):e51 [FREE Full text] [doi: 10.2196/jmir.2814] [Medline: 24550161]
24. Pinnock H. Supported self-management for asthma. Breathe (Sheff) 2015 Jun 03;11(2):98-109 [FREE Full text] [doi: 10.1183/20734735.015614] [Medline: 26306110]
25. James DR, Lyttle MD. British guideline on the management of asthma: SIGN Clinical Guideline 141, 2014. Arch Dis Child Educ Pract Ed 2016 Dec 20;101(6):319-322. [doi: 10.1136/archdischild-2015-310145] [Medline: 27206454]
26. Jabangwe R, Edison H, Duc AN. Software engineering process models for mobile app development: a systematic literature review. J Syst Softw 2018 Nov;145:98-111. [doi: 10.1016/j.jss.2018.08.028]
27. van Gemert-Pijnen JE, Nijland N, van Limburg M, Ossebaard HC, Kelders SM, Eysenbach G, et al. A holistic framework to improve the uptake and impact of eHealth technologies. J Med Internet Res 2011 Dec 05;13(4):e111 [FREE Full text] [doi: 10.2196/jmir.1672] [Medline: 22155738]
28. van den Berg LN, Hallensleben C, Chavannes NH, Versluis A. Developing a smartphone application that promotes responsible short-acting beta2-agonist use in people with asthma: a participatory design. Int J Environ Res Public Health 2022 Jul 12;19(14):8496 [FREE Full text] [doi: 10.3390/ijerph19148496] [Medline: 35886348]
29. Guest G, Bunce A, Johnson L. How many interviews are enough? An experiment with data saturation and variability. Field Methods 2016 Jul 21;18(1):59-82. [doi: 10.1177/1525822X05279903]
30. Guest G, Namey E, Chen M. A simple method to assess and report thematic saturation in qualitative research. PLoS One 2020 May 5;15(5):e0232076 [FREE Full text] [doi: 10.1371/journal.pone.0232076] [Medline: 32369511]

XSL•FO
RenderX

31. Ware JE, Sherbourne CD. The MOS 36-item short-form health survey (SF-36). I. Conceptual framework and item selection. Med Care 1992 Jun;30(6):473-483. [Medline: 1593914]
32. Huijg JM, Gebhardt WA, Crone MR, Dusseldorp E, Presseau J. Discriminant content validity of a theoretical domains framework questionnaire for use in implementation research. Implement Sci 2014 Jan 15;9(1):11 [FREE Full text] [doi: 10.1186/1748-5908-9-11] [Medline: 24423394]
33. Azevedo P, Correia de Sousa J, Bousquet J, Bugalho-Almeida A, Del Giacco SR, Demoly P, et al. Control of Allergic Rhinitis and Asthma Test (CARAT): dissemination and applications in primary care. Prim Care Respir J 2013 Mar 14;22(1):112-116 [FREE Full text] [doi: 10.4104/pcrj.2013.00012] [Medline: 23412110]
34. The control of allergic rhinitis and asthma test. The Control of Allergic Rhinitis and Asthma Test Network. URL: https://www.new.caratnetwork.org/fastcarat/index.html [accessed 2021-05-24]
35. Brooke J. SUS: a 'quick and dirty' usability scale. In: Usability Evaluation In Industry. Boca Raton, FL: CRC Press; 1996.
36. Castor EDC. URL: https://www.castoredc.com/ [accessed 2020-02-17]
37. Venkatesh V, Morris MG, Davis GB, Davis FD. User acceptance of information technology: toward a unified view. MIS Q 2003 Sep;27(3):425-478. [doi: 10.2307/30036540]
38. McHorney CA, Ware JE, Raczek AE. The MOS 36-item short-form health survey (SF-36): II. Psychometric and clinical tests of validity in measuring physical and mental health constructs. Med Care 1993 Mar;31(3):247-263. [doi: 10.1097/00005650-199303000-00006] [Medline: 8450681]
39. Ware JE, Kosinski M, Keller SD. SF-36 Physical and Mental Health Summary Scales: A User's Manual. Boston, MA: Health Assessment Lab; 1994.
40. Grassi M, Nucera A, European Community Respiratory Health Study Quality of Life Working Group. Dimensionality and summary measures of the SF-36 v1.6: comparison of scale- and item-based approach across ECRHS II adults population. Value Health 2010;13(4):469-478 [FREE Full text] [doi: 10.1111/j.1524-4733.2009.00684.x] [Medline: 20088893]
41. IBM SPSS Statistics 25.0. IBM. URL: https://www.ibm.com/support/pages/downloading-ibm-spss-statistics-25 [accessed 2020-03-24]
42. Gale NK, Heath G, Cameron E, Rashid S, Redwood S. Using the framework method for the analysis of qualitative data in multi-disciplinary health research. BMC Med Res Methodol 2013 Sep 18;13:117 [FREE Full text] [doi: 10.1186/1471-2288-13-117] [Medline: 24047204]
43. ATLAS.ti homepage. ATLAS.ti Scientific Software Development GmbH. URL: https://atlasti.com [accessed 2022-01-13]
44. van Buul AR, Kasteleyn MJ, Arends JM, Shi T, Kelly DP, Chavannes NH, et al. eHealth only interventions and blended interventions to support self-management in adolescents with asthma: a systematic review. Clinical eHealth 2020;3:49-62. [doi: 10.1016/j.ceh.2020.06.001]
45. Gonzalez-Barcala FJ, de la Fuente-Cid R, Tafalla M, Nuevo J, Caamaño-Isorna F. Factors associated with health-related quality of life in adults with asthma. A cross-sectional study. Multidiscip Respir Med 2012 Oct 02;7(1):32 [FREE Full text] [doi: 10.1186/2049-6958-7-32] [Medline: 23031194]
46. Chen H, Gould MK, Blanc PD, Miller DP, Kamath TV, Lee JH, et al. Asthma control, severity, and quality of life: quantifying the effect of uncontrolled disease. J Allergy Clin Immunol 2007 Aug;120(2):396-402. [doi: 10.1016/j.jaci.2007.04.040] [Medline: 17561244]
47. Hui CY, Walton R, McKinstry B, Jackson T, Parker R, Pinnock H. The use of mobile applications to support self-management for people with asthma: a systematic review of controlled studies to identify features associated with clinical effectiveness and adherence. J Am Med Inform Assoc 2017 May 01;24(3):619-632 [FREE Full text] [doi: 10.1093/jamia/ocw143] [Medline: 27694279]
48. Beller J, Geyer S, Epping J. Health and study dropout: health aspects differentially predict attrition. BMC Med Res Methodol 2022 Jan 30;22(1):31 [FREE Full text] [doi: 10.1186/s12874-022-01508-w] [Medline: 35094681]
49. Sulaiman I, Greene G, MacHale E, Seheult J, Mokoka M, D'Arcy S, et al. A randomised clinical trial of feedback on inhaler adherence and technique in patients with severe uncontrolled asthma. Eur Respir J 2018 Jan 04;51(1):1701126 [FREE Full text] [doi: 10.1183/13993003.01126-2017] [Medline: 29301919]
50. van de Hei SJ, Poot CC, van den Berg LN, Meijer E, van Boven JF, Flokstra-de Blok BM, et al. Effectiveness, usability and acceptability of a smart inhaler programme in patients with asthma: protocol of the multicentre, pragmatic, open-label, cluster randomised controlled ACCEPTANCE trial. BMJ Open Respir Res 2022 Dec;9(1):e001400 [FREE Full text] [doi: 10.1136/bmjresp-2022-001400] [Medline: 36522130]
51. Renzi E, Baccolini V, Migliara G, De Vito C, Gasperini G, Cianciulli A, et al. The impact of eHealth interventions on the improvement of self-care in chronic patients: an overview of systematic reviews. Life (Basel) 2022 Aug 17;12(8):1253 [FREE Full text] [doi: 10.3390/life12081253] [Medline: 36013432]
52. van der Kleij RM, Kasteleyn MJ, Meijer E, Bonten TN, Houwink EJ, Teichert M, et al. SERIES: eHealth in primary care. Part 1: concepts, conditions and challenges. Eur J Gen Pract 2019 Oct;25(4):179-189 [FREE Full text] [doi: 10.1080/13814788.2019.1658190] [Medline: 31597502]
53. Song X, Hallensleben C, Zhang W, Jiang Z, Shen H, Gobbens RJ, et al. Blended self-management interventions to reduce disease burden in patients with chronic obstructive pulmonary disease and asthma: systematic review and meta-analysis. J Med Internet Res 2021 Mar 31;23(3):e24602 [FREE Full text] [doi: 10.2196/24602] [Medline: 33788700]

XSL•FO
RenderX

54. Fixsen DL, Blasé KA, Timbers GD, Wolf MM. In search of program implementation: 792 replications of the teaching-family model. Behav Anal Today 2007;8(1):96-110. [doi: 10.1037/h0100104]
55. Versluis A, van Luenen S, Meijer E, Honkoop PJ, Pinnock H, Mohr DC, et al. SERIES: eHealth in primary care. Part 4: addressing the challenges of implementation. Eur J Gen Pract 2020 Dec;26(1):140-145 [FREE Full text] [doi: 10.1080/13814788.2020.1826431] [Medline: 33025820]
56. Khalid A, Dong Q, Chuluunbaatar E, Haldane V, Durrani H, Wei X. Implementation science perspectives on implementing telemedicine interventions for hypertension or diabetes management: scoping review. J Med Internet Res 2023 Mar 14;25:e42134 [FREE Full text] [doi: 10.2196/42134] [Medline: 36917174]
57. Ross J, Stevenson F, Lau R, Murray E. Factors that influence the implementation of e-health: a systematic review of systematic reviews (an update). Implement Sci 2016 Oct 26;11(1):146 [FREE Full text] [doi: 10.1186/s13012-016-0510-7] [Medline: 27782832]
58. Tossaint-Schoenmakers R, Versluis A, Chavannes N, Talboom-Kamp E, Kasteleyn M. The challenge of integrating eHealth into health care: systematic literature review of the Donabedian model of structure, process, and outcome. J Med Internet Res 2021 May 10;23(5):e27180 [FREE Full text] [doi: 10.2196/27180] [Medline: 33970123]

Abbreviations

CARAT: Control of Allergic Rhinitis and Asthma Test
ICS: inhaled corticosteroids
QoL: quality of life
SABA: short-acting beta2-agonist
SF-36: 36-Item Short-Form Health Survey
SUS: System Usability Scale
UTAUT: Unified Theory of Acceptance and Use of Technology

Edited by G Eysenbach, A Kushniruk; submitted 08.11.23; peer-reviewed by UK Chalwadi, D Drummond; comments to author 06.01.24; revised version received 26.01.24; accepted 18.02.24; published 04.04.24.

Please cite as:
van den Berg LN, Hallensleben C, Vlug LAE, Chavannes NH, Versluis A
The Asthma App as a New Way to Promote Responsible Short-Acting Beta2-Agonist Use in People With Asthma: Results of a Mixed Methods Pilot Study
JMIR Hum Factors 2024;11:e54386
URL: https://humanfactors.jmir.org/2024/1/e54386
doi:10.2196/54386
PMID:38574348

©Liselot N van den Berg, Cynthia Hallensleben, Lisa AE Vlug, Niels H Chavannes, Anke Versluis. Originally published in JMIR Human Factors (https://humanfactors.jmir.org), 04.04.2024. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on https://humanfactors.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

A Smartphone App to Support Self-Management for People Living With Sjögren's Syndrome: Qualitative Co-Design Workshops

Claire McCallum[1*], PhD; Miglena Campbell[2*], PhD; John Vines[3], PhD; Tim Rapley[4], PhD; Jason Ellis[5], PhD; Vincent Deary[5], PhD; Katie Hackett[4], PhD

[1]Faculty of Engineering, University of Bristol, Bristol, United Kingdom
[2]Institute for Collective Place Leadership, Teesside University, Middlesbrough, United Kingdom
[3]School of Informatics, University of Edinburgh, Edinburgh, United Kingdom
[4]Department of Social Work, Education and Community Wellbeing, Northumbria University, Newcastle upon Tyne, United Kingdom
[5]Department of Psychology, Northumbria University, Newcastle upon Tyne, United Kingdom
[*]these authors contributed equally

Corresponding Author:
Katie Hackett, PhD
Department of Social Work, Education and Community Wellbeing
Northumbria University
Coach Lane Campus West
Newcastle upon Tyne, NE7 7XA
United Kingdom
Phone: 44 191 215 6098
Email: kate.hackett@northumbria.ac.uk

Abstract

Background: Sjögren's syndrome (SS) is the second most common autoimmune rheumatic disease, and the range of symptoms includes fatigue, dryness, sleep disturbances, and pain. Smartphone apps may help deliver a variety of cognitive and behavioral techniques to support self-management in SS. However, app-based interventions must be carefully designed to promote engagement and motivate behavior change.

Objective: We aimed to explore self-management approaches and challenges experienced by people living with SS and produce a corresponding set of design recommendations that inform the design of an engaging, motivating, and evidence-based self-management app for those living with SS.

Methods: We conducted a series of 8 co-design workshops and an additional 3 interviews with participants who were unable to attend a workshop. These were audio recorded, transcribed, and initially thematically analyzed using an inductive approach. Then, the themes were mapped to the Self-Determination Theory domains of competency, autonomy, and relatedness.

Results: Participants experienced a considerable demand in the daily work required in self-managing their SS. The condition demanded unrelenting, fluctuating, and unpredictable mental, physical, and social efforts. Participants used a wide variety of techniques to self-manage their symptoms; however, their sense of competency was undermined by the complexity and interconnected nature of their symptoms and affected by interactions with others. The daily contexts in which this labor was occurring revealed ample opportunities to use digital health aids. The lived experience of participants showed that the constructs of competency, autonomy, and relatedness existed in a complex equilibrium with each other. Sometimes, they were disrupted by tensions, whereas on other occasions, they worked together harmoniously.

Conclusions: An SS self-management app needs to recognize the complexity and overlap of symptoms and the complexities of managing the condition in daily life. Identifying techniques that target several symptoms simultaneously may prevent users from becoming overwhelmed. Including techniques that support assertiveness and communication with others about the condition, its symptoms, and users' limitations may support users in their interactions with others and improve engagement in symptom management strategies. For digital health aids (such as self-management apps) to provide meaningful support, they should be designed according to human needs such as competence, autonomy, and relatedness. However, the complexities among the 3 Self-Determination Theory constructs should be carefully considered, as they present both design difficulties and opportunities.

(JMIR Hum Factors 2024;11:e54172) doi:10.2196/54172

XSL•FO
RenderX

KEYWORDS

self-management; mobile health; mHealth; eHealth; Sjögren's syndrome; patient participation; patient involvement; fatigue; chronic disease; focus groups; complex intervention development; mobile phone

Introduction

Background

The need to improve the accessibility and quality of care for those with long-term conditions (LTCs) is an international priority [1]. In England alone, LTCs affect 15 million people [2] and account for 70% of health care spending [3]. Rheumatic diseases are LTCs with a particularly high prevalence in the United Kingdom and worldwide, having been estimated to affect up to one-fourth of Europeans [4,5] and a similar proportion of the population in the global south [6]. Sjögren's syndrome (SS) is thought to be the second most common autoimmune rheumatic disease [7] and is associated with poor quality of life [8] and high disease burden [9].

SS is a heterogeneous LTC, with a constellation of unpredictable and diverse symptoms [10,11]. A key characteristic of SS is mucosal dryness due to the destruction of exocrine (moisture-producing) glands by the body's immune system, which particularly affects the eyes, mouth, and vagina [12]. In addition to dryness, common extraglandular features include persistent fatigue [13], chronic pain [14], sleep disturbances [15], and anxiety and depression [16]. People with SS report experiencing these symptoms as being interconnected, with the exacerbation of one symptom impacting others [17-19].

Similar to many other autoimmune diseases, SS does not have a cure [20]. Therefore, intervention efforts have focused on reducing the severity of symptoms; for instance, topical treatments are used for managing dryness [21]. Drug treatments for the systemic management of SS, such as hydroxychloroquine and rituximab, have had disappointing results in clinical trials [22,23]. Behavioral interventions that aim to improve the quality of life are a promising alternative; however, few interventions have been developed, and evaluations of their impact have been of low quality [21,24]. A recent stakeholder engagement study found that support for self-managing symptoms was a key priority for people with SS [25]. The term self-management has been defined as "the individual's ability to manage the symptoms, treatment, physical and psychosocial consequences and lifestyle changes inherent in living with a chronic condition" [26,27]. To support the knowledge, behaviors, and attitudes required, self-management interventions should deliver a range of educational, behavioral, and cognitive techniques [28]. In SS, a targeted "complex" intervention is required, which delivers multiple techniques and targets multiple SS symptoms [29]. Our previous body of work with patients with SS found that they require different levels of support. Some require more complex individual support, but most people require lower levels of support with access to written information and digital self-management tools [29], which could be provided in the form of a website or smartphone app.

Apps as a Support for Self-Management

SS shares multiple symptom and self-management similarities with other LTCs [30], including but not limited to neurological and autoimmune conditions such as rheumatoid arthritis, myalgic encephalomyelitis, and multiple sclerosis. Smartphone apps are a promising approach to support self-management of these LTCs [31,32] and other conditions such as type 2 diabetes [33], asthma [34], and hypertension [35]. Their increasing availability and functionalities enable complex intervention techniques to be delivered in the context of users' daily lives when they are designed with consideration of users' routines and choices [36]. User-centered design studies of LTCs have produced various app features and content [37] to support, for example, user education and cognitive strategies. However, app effectiveness can be limited by very low levels of user engagement [38,39]. Therefore, intervention developers must design apps that are more engaging and carefully consider how such engagement will ultimately lead to long-term behavior change [40]. For example, beyond simply providing information about how to perform techniques, apps can be designed to promote a sense of autonomy and motivation to engage in self-management behaviors over time [41,42].

To increase user engagement, apps should be user centered and person centered [43], that is, designed to fit within individuals' current lives and daily activities [44]. People are more likely to use a new intervention if it can be incorporated into their existing habits, routines, and contexts [45,46]. Therefore, self-management interventions should account for and actively support how people manage their conditions currently [31,47]. Thus, to develop a useful, effective, and engaging app-based intervention that supports those with SS, there is a need to first understand their current self-management opportunities and challenges. To date, limited studies have only been conducted to understand the lived experience of symptoms [17,19,48] and have not explored the self-management of multiple diverse symptoms.

To gain an understanding of individuals' self-management contexts, co-design and user-centered methods are useful [49]. These can involve practical design activities that elicit conversations regarding a topic of interest (such as self-management) to inform the development of a design, product, or intervention and have been used to develop digital health interventions [50,51]. Then, to understand how users in these contexts might best be supported in changing their self-management behavior, co-design findings can be interpreted using theories of motivation and behavior change [41].

Self-Determination Theory (SDT) [52] is one motivational theory widely used in interventions promoting health behavior change [53,54], including those for self-managing chronic illnesses [55,56]. SDT proposes that the constructs of competence, autonomy, and relatedness are required for individuals to be internally motivated to perform behaviors and sustain these changes over time. Situating qualitative findings

XSL•FO
RenderX

within theoretical constructs facilitates the development of apps that are based on theory [42,57]. While intervention developers use SDT to inform their interventions, many do not explicitly link the theoretical constructs directly to their individual components, and we aimed to bridge this gap. To the best of our knowledge, there are no evidence-based, theory-driven, self-management apps for SS.

Study Aims

We aimed to use an SDT framework to explore self-management challenges and approaches used by people with SS and to produce a set of design and therapeutic recommendations for a supportive and engaging app to aid self-management.

Methods

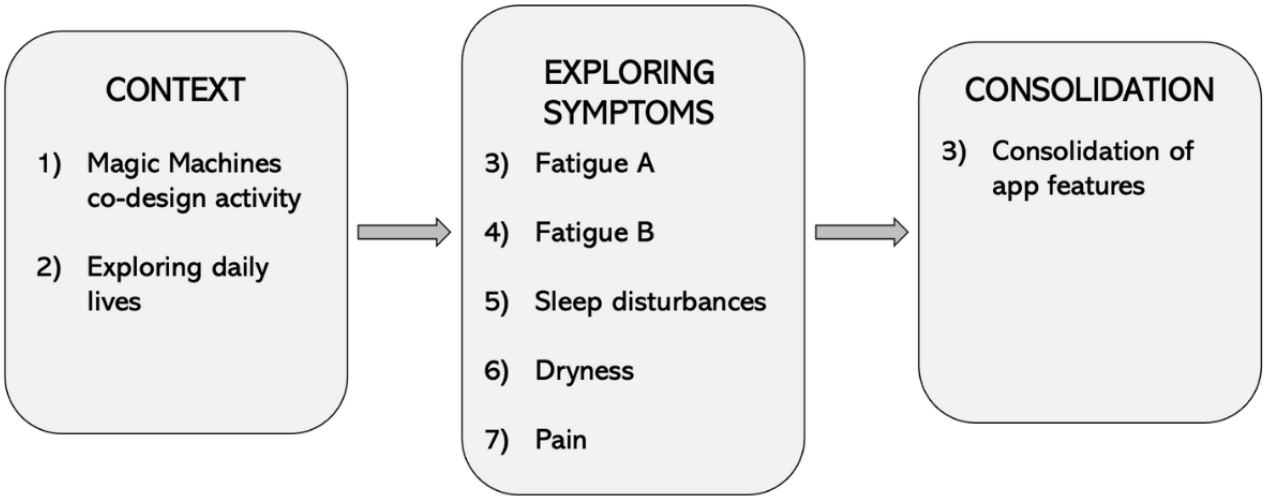
The methods and subsequent results have been reported according to the COREQ (Consolidated Criteria for Reporting Qualitative Research) guidelines (Multimedia Appendix 1 [58]).

Study Design

A consecutive series of 8 workshops with people living with SS was conducted over 7 months, each involving design activities and focused discussions (Figure 1). The first 2 workshops were open ended to broadly understand participants' contexts (ie, key self-management challenges and overall self-management routines) and enable participants to become familiar with each other and to feel comfortable while discussing potentially sensitive and personal topics. We decided in advance to include a series of workshops, with each workshop dedicated to in-depth understanding of the self-management activities, challenges, and opportunities for each symptom. However, the order of symptom workshops and their exact discussion topics and activities were not predetermined; their sequential nature enabled us to iteratively design topics based on the findings from the previous session. For example, a clear theme emerged regarding symptom interrelatedness, so subsequent workshops included discussions about how participants managed interrelations among their symptoms. Furthermore, fatigue was a priority for all workshop participants (14/14, 100%), and therefore, 2 workshops were dedicated to this symptom.

Participants were given the option to attend ≥1 workshops. Several workshops were repeated to suit participants' availability. To enable those who could not access any workshops due to other commitments and to include the experiences of younger people living with SS, 3 one-off semistructured interviews were conducted. These focused on the key self-management practices and challenges experienced by the respective participant.

Figure 1. The procedural flow and topics of the 8 design workshops.



Ethical Considerations

This qualitative study received ethics approval from Northumbria University ethics committee (reference 11130). Informed consent was obtained before data collection, and travel costs were reimbursed.

Recruitment

Workshop participants were purposively recruited from a regional UK SS support group (Northeast Sjögren's Syndrome Association). Advertisements were distributed via their member mailing list and Facebook page, and the research team presented the project at a support group meeting. The invitation was open to those diagnosed with SS by a physician, and potential participants were invited to attend as many workshops as they liked. Interested participants who were unable to attend due to their location or life commitments were invited to attend a video web-based interview. Additional participants were recruited via social media (a single tweet on Twitter [subsequently rebranded as X]) and invited to participate in the interviews only.

Data Collection Activities

Overview

Workshops were conducted at Northumbria University, lasted approximately 90 minutes, and included a 10-minute comfort break. The interviews lasted 30 to 60 minutes and were conducted via telephone or videoconferencing software. Workshops were facilitated by 3 authors (CM, MC, and KH)—all were female postdoctoral (PhD) researchers trained in qualitative research methods and experienced in conducting qualitative research interviews and focus groups; one of them

XSL•FO
RenderX

was also an occupational therapist (KH) with experience in SS symptom management. Several workshop participants (4/14, 29%) had attended clinics (conducted by KH), and 14% (2/14) had participated in previous studies (conducted by KH and VD). All participants (17/17, 100%) were briefed about the aims of the study. All workshops and interviews were audio recorded, and facilitators took field notes. In the following sections, we have outlined the focus of each workshop. The individual workshop topic guides are presented in Multimedia Appendix 2 [59-62].

Workshop 1: Magic Machines Co-Design Activity

This workshop introduced the series of workshops, included discussion about some key self-management issues experienced by participants, and involved a Magic Machines [54,59] craft activity where participants created some imaginative design solutions for another workshop participant. The Magic Machines activity aimed to elicit a broad range of knowledge about participants' personal and technological needs through discussions about everyday problems related to their condition and potential solutions. Participants were asked to create an object, which addressed their partner's daily challenge, using household objects and craft items. Data capture was focused on the conversations between participants about their "problem" while making their objects (a potential "solution") and when describing their object to the main group at the end of the session.

Workshop 2: Exploring Daily Lives

The second workshop explored individuals' "daily lives" and the self-management of symptoms. The discussion about daily lives invited participants to discuss their "typical day" in managing SS (ie, their habits and routines), how SS self-management was incorporated into their routines, and any related challenges that they experienced.

Workshops 3 to 7: Exploring Symptoms

These workshops explored the self-management of specific symptoms and their interrelationships through group discussions and invited participants to engage in basic sketching to articulate their self-management experiences and challenges. We preselected the symptoms for discussion based on our previous study where patients identified them as being important and impacting their daily activities [25].

Workshop 8: Consolidation

Sketching was used to explore how an app might be structured to support symptom interconnectedness and complexity. This design activity also elicited discussion about user experience and usability issues. All participants (5/5, 100%) attending this workshop engaged in sketching, but if time was insufficient, they were encouraged to further develop their ideas by articulating them verbally.

Interviews

Following the workshops, 3 semistructured, web-based interviews were conducted by CM. The interviews followed a schedule of open-ended questions to allow for flexibility (Multimedia Appendix 3).

Data Analysis

Audio data were transcribed verbatim, pseudonymized, and combined into a corpus for analysis using NVivo (version 12; QSR International). Analysis was conducted in 2 phases using a hybrid approach, incorporating both inductive and deductive methodologies, to harness the advantages of both methods [63]. First, an inductive thematic analysis approach [64] was used, where 2 researchers (CM and MC) independently coded the data, generating an initial set of codes related to participants' self-management perceptions and experiences. Then, these codes were applied and refined through the arrival of each new transcript, and independent coding was subsequently conducted by CM. Discussions during regular research team meetings (with CM, MC, and KH) related to the codes and their connections, importance, and relevance were conducted to group codes into themes.

Then, these inductive themes were mapped to the 3 SDT [65] constructs of competency (the sense of capability to perform activities and tasks), autonomy (experience of having control and choice over one's actions and decisions), and relatedness (feeling of connection and belonging and meaningful interaction with others) by CM. SDT was chosen over other motivational theories because it emphasizes social context as a key factor in helping or hindering motivation, which matched a prominent theme in our inductive thematic analysis of social relations, along with other major themes we found related to empowerment, autonomy, and capability (or "competency" in SDT). The theory is also highly translational, enabling findings to inform intervention design [66]. Regular research team meetings were conducted to review and reach consensus regarding the categorization of themes based on the SDT constructs. Opportunities to support participants' challenges associated with these themes through an app were also identified through discussions. Methodological rigor and credibility of findings were pursued through development of a codebook, maintenance of ongoing reflexivity, peer debriefings, and data triangulation (from interviews, focus groups, and observations during workshop activities).

Results

Participants

In total, 17 people with SS participated in the workshops and interviews: 14 (82%) of the 17 participated in the workshops (13/14, 93% women and 1/14, 7% men) and 3 (18%) of the 17 participated in the web-based interviews (3/3, 100% women). Participants' ages ranged from 33 to 76 (mean 56.5, SD 13.95) years, and 82% (14/17) of them had a diagnosis of primary SS. The remaining 18% (3/17) of the participants had a diagnosis of secondary SS. The mean number of years since diagnosis was 7.5 (SD 7.88) years. Regarding employment status, of the 17 participants, 8 (47%) were retired, 6 (35%) were working full time, 1 (6%) was in part-time employment, and 2 (12%) were not working currently. All workshop participants (14/14, 100%) had links to a local SS support group in the north of England. Of the 3 interviewees, 1 (33%) was part of the same support group, and the remaining 2 (67%) were from Spain and Canada, respectively (both were aged <35 years). Workshop

XSL•FO
RenderX

group numbers ranged from 2 to 7 participants and 2 to 3 facilitators in each session. Some participants attended some sessions and not others, whereas a "core" group of 5 participants attended most sessions. One participant (1/14, 7%) attended only the first workshop; others attended at least 2 sessions.

Overview

Participants engaged in a wide range of self-management behaviors, including using prescribed and over-the-counter medications and treatments (ie, applying eye drops and gels; bathing and massaging the eyes; using humidifiers, skin creams, and vaginal lubricants; following mouth care routines; using pain medication; and using hot and cold compresses). They also used cognitive and behavioral techniques including activity pacing, goal setting, general exercise, relaxation, mindfulness, distraction, napping, sleep management and wind-down routines, and social support. Participants used various tools to support and facilitate the learning, use, and practice of these techniques, including books (eg, about managing fatigue), diaries (paper based and digital), websites and forums (eg, National Health Service or SS associations as both knowledge resources and social support), apps on smartphones and tablets (such as for yoga, breathing exercises, and mindfulness), wearables (to track physical activity), and other devices (eg, for relaxing music, "mindless" television, or distracting podcasts and comedy). Not all participants owned or used smartphones. Tools were used in addition to visiting friends and holistic wellness centers (eg, spas and mindfulness classes) and learning self-care techniques directly from health care professionals (eg, when to apply eye drops and more complex techniques such as activity pacing and graded exercise).

In the following sections, we have described the challenges that participants faced in managing their condition and their psychosocial needs in terms of competency, autonomy, and relatedness.

Competency

Participants varied in the extent to which they felt competent and successful in self-managing their SS, and this was related to how well they had established a self-care routine. One participant had a very "strict regime," which they felt was required to maintain their level of functioning. While hearing about such self-management strategies from others, Jim reflected about his competencies:

> *I've still got quite a lot to learn...although it has been a few years now, I think I still haven't got a good routine...I listen to your explanation [of another participant's routine] and I think, why can't I get myself like that? I'm supposed to be Mr Organised. I am known as that in my life. My working life and my own home life. Yet with this, I have not gotten organised yet.* [Jim]

Regardless of whether participants had routines or described habitual self-management behaviors, their sense of competency in self-managing SS was still impeded by the complex nature of their symptoms. Isolating and targeting individual symptoms was not only perceived to be difficult to perform ("You can't separate the different symptoms" [Jim]) but were also sometimes unhelpful, as it did not account for their accumulative negative impact:

> *It is the overall effect to me. That three [symptoms] I can cope with and then the next day one raises its head and floors me...That straw that broke the camel's back effect, you know. [Patricia]*

Several participants believed that they could better manage their symptoms through self-management techniques capable of improving multiple symptoms simultaneously. Some had discovered these types of techniques accidentally. For example, participants recounted noticing, with surprise, that eye drops had helped not only their dryness but also mental and physical fatigue. Other participants purposefully sought and regularly used techniques that targeted multiple symptoms simultaneously. Mindfulness and relaxation techniques provided a sense of control and the ability to "keep a cap on" multiple symptoms before they became very severe. Others agreed that seeking these techniques was worthwhile if they resulted in minor improvements across multiple symptoms. Despite valuing self-management techniques that targeted multiple symptoms, most did not feel confident or knowledgeable about which techniques were beneficial.

Another challenge to participants' sense of competency is how SS symptoms are not static but change over time. Participants described instances where individual symptoms would rapidly fluctuate in severity:

> *[They] come and go...one day you might have a headache, the next day you don't. [Jim]*

Participants also explained experiencing longer periods where multiple symptoms were severe (described using phrases such as a "flare," "phase," or "wave") or individual symptoms persisted (such as "a dry patch"). While, sometimes, the onset of symptoms appeared "gradually," at other times, they changed rapidly, leaving participants feeling unprepared ("a phase hits you"). Fatigue and pain were felt to be particularly volatile and could become severe with no warning and "like somebody just switched a switch" (Penny).

Participants varied in how they managed such changing symptoms. Many attended to symptoms as they arose or increased in severity on a moment-to-moment basis (ie, an adaptive or reactive approach). However, this often meant devising complex and intricate strategies and sequences to manage the new combination of symptoms experienced in that moment. For example, sleep disturbances that might be attributed to pain, dryness, or anxiety required participants to change their approach to getting back to sleep accordingly ("depending on how I am" [Penny] or "what problem I am having" [Jim]). Other participants seemed to disregard the changing combination of symptoms and addressed symptoms "one at a time" based on whether they felt successfully managed. For example, Julie noted the following:

> *I tend to find like I feel like my feet are sorted, so I am now sorting my eyes, so I'm kind of going through this list.*

Addressing symptoms required constant adjustments for participants. Their variable nature meant that just at the point

XSL•FO
RenderX

that the individual starts to feel in control of one symptom, a flare of another may occur.

A final layer of complexity impacting participant competency was how symptoms often change due to environmental factors. For example, dryness was exacerbated by air conditioning, bright lighting, and other people's aftershaves and perfumes, whereas navigating new and busy places could exacerbate mental fatigue. The unpredictable nature of environments outside the home made self-managing symptoms more challenging. While home was characterized as "familiar" and "unchanging," participants felt that they needed to continuously estimate the potential impact of environments on their symptoms and plan accordingly:

> *You have to be very wary of where you're going...you've got to be careful. I will not walk through [the shopping mall] in the perfumery because there is always somebody going to...pick up a bottle of perfume and [spray]...I go, oh my eyes! [Geraldine]*

This planning itself was exhausting to several participants, and it also meant that they lacked spontaneity in their lives. Participants also felt that symptoms were easier to manage at home because they could easily perform physical relaxation and self-care techniques when required, particularly during a flare. During such times Sarah remarked as follows:

> *I just don't want to leave the house, I don't want to do anything. I just want to go and have like 2-3 baths per day.*

In contrast, when symptoms left them debilitated outside the house, participants had to adopt different self-management techniques such as soothing self-talk or be "rescued" by a taxi or friend. Overall, being outside the home meant that participants were less in control of their environments; had to continuously plan and predict how the environment might impact their symptoms, which was mentally and physically tiring; and had to use different techniques to suit different environments.

Autonomy

Our analysis identified many examples of participants feeling that they had autonomy in the self-management of their condition; however, sometimes, the same factors that promoted autonomy also reduced confidence and competence. Participants believed that the availability of various techniques meant that they had options in their self-management; there were multiple "different ways" they could try to improve symptoms. The plurality of techniques appeared to provide reassurance that at least one would be likely to be effective:

> *I have six choices...I don't beat myself up when it doesn't work because I've already got something else in mind.* [Patricia]

This plurality and optimism could provide a strong drive to continue in their self-management activities.

Participants varied in how they kept track of different available techniques. One participant had self-help books at various locations in their home. Another participant explained that they had collated several techniques to create their own book:

> *I wrote myself a little book...[of] top tips...I just wrote maybe two dozen messages across the book at random, things that might give me a clue. [Debra]*

Other participants used an experimental approach:

> *It's about learning...through trial and error...you'll notice a pattern...you don't know until you've done it for a few months.* [Michelle]

These were similarly characterized by the desire to try different techniques and to keep track of their effectiveness:

> *With time and experience you begin to realise what works and what doesn't.* [Penny]

It was acknowledged that this required continuous effort and perseverance.

Having personal choice to decide which techniques to try, as opposed to being directed by a health care professional, provided some participants with a sense of control. Debra likened creating her book of techniques to developing a tailored smartphone app:

> *It is basically my own app that I've written for myself...I didn't feel like being ordered around by anybody else...I don't necessarily follow it. If it's inside my book, I think, well alright, maybe I'll try something else...I've still got some kind of control over things. [Debra]*

Therefore, developing this herself meant that she did not feel obligated to try any 1 technique. Although participants appreciated having the autonomy to choose techniques in a personalized manner, the credibility of these techniques was also very important to them. Perceived credibility seemed to give them confidence to go ahead and try them. Some participants indicated that they understood the distinction between evidence-based information and hearsay:

> *I am pretty much someone who will try anything once if there's some evidence to support its effectiveness...Some people suggest real outlandish things, like you hear it and you're like, "okay!" I mean, I'm glad that it works for you, but I'm not really sold on trying that just yet.* [Ellie]

Participants felt that information about their condition or how to manage symptoms should be credible. For example, Jim explained that simply being presented with multiple self-management techniques and options, without a rationale for why they might be helpful, would not suit him. Others stated that knowing information sources was "useful...[for] controlling symptoms and trying to minimise [them]" (Edith). Information from websites such as the UK National Health Service or regional and national SS organizations was deemed trustworthy.

Although participants respected expert advice and implemented it in their self-management, expert authority was often only 1 element in an autonomous process of symptom management decision-making. For instance, when faced with a conflict between their preferred routines and expert advice, participants trusted their own expertise and experience. Jim outlined how he fell asleep with the help of music or old comedy shows and that he would simply "ignore" any potential prompts about

XSL•FO
RenderX

adjusting his bedtime routine if it meant removing his music from the bed (as may be advised as part of a sleep intervention).

For some participants, smartphones and associated apps appeared to contribute to feelings of autonomy regarding their self-management of SS. Those who used a smartphone reported using basic note apps to track symptoms or calendars built into the operating system to track feelings of fatigue. Experienced smartphone users described how their ubiquitous nature enabled quick access to information and could give them access to techniques whenever and wherever needed, regardless of their location. In sessions where feedback was given about potential app designs, participants expressed the value they would see in new apps that brought various techniques together, provided reminders to apply eye drops, and helped track symptoms in a simpler manner. For instance, Julie suggested that "a tracker or a journal...or something like that on the app would be helpful" as this could help her manage her forgetfulness, which she referred to as "brain fog."

However, while smartphones could enable autonomy, they also posed challenges that could impact the users' SS symptoms. It was also noted that looking at the screen of a computer or smartphone for a very long time could exacerbate eye dryness:

> *It's okay [when it's] short, but you can't spend a long time looking at the screen, because your eyes are just too sore.* [Mel]

To overcome this, participants used their smartphones differently. Some described deliberately limiting the amount of time they used them in 1 session, and others described changing their device settings to increase the font size or darken the screens. In addition, participants mentioned improving on-screen accessibility to reduce their eye strain and listening to audio instead of reading text. Patricia, noted that when "I am having my brain fog" the complexity of most apps "would blow my mind." Among participants, there was a sense that smartphone use was closely related to experiences of mental fatigue from their SS.

Overall, participants valued the diversity of SS self-management techniques that are available and experienced this as enhancing autonomy. Smartphones and both generic and SS-specific apps were viewed as an important part of this diversity and could provide in situ tailored support. However, the apparent abundance of techniques and availability of smartphones also posed a challenge to autonomy. Patricia recounted that soon after being diagnosed with SS, she was overwhelmed by the need to learn about multiple symptoms and techniques from many sources. However, for her, this felt similar to being "shot at" from multiple angles. Sometimes, the factors that enable autonomy can also constrain it.

Relatedness

Relatedness refers to the manner in which participants operated in their social worlds and how their practices of managing SS were related to it. Participants explained that SS profoundly impacted their familial interactions, friendships, and other forms of social contact. Participants enjoyed social activities and cherished positive relationships as a source of social support. Socializing and participating in activities with others provided a positive "distraction" from their symptoms. However, self-management tasks could impact their ability to socialize and interrupt the flow of conversations:

> *When in company if you are out and about and talking to people...You have to keep popping off to go and put eyedrops in, in the loo.* [Edith]

Furthermore, engaging in certain social activities, such as going to the cinema with friends, required participants to perform additional self-care, for example, applying eye drops more frequently, which could irritate the skin around the eyes. Geraldine explained that although she enjoyed going to live theatre performances, she was now reluctant to go based on previously being "crucified" by a smoke machine.

Pacing was a helpful technique to manage fatigue, but it was not always received well by others in social situations and workplaces. Patricia recounted that she had been regarded as "selfish" by family and friends for cancelling plans while trying to manage her energy and fatigue levels. She also recognized that having to "book" people into her diary well in advance to support her planning and pacing efforts "frightens some people off." Edith recalled that the need to take more breaks meant that she had to decide to leave her walking group as she was no longer able to keep up with her friends. In turn, this negatively impacted her feeling of belonging.

Communication was key while managing illness demands and relationships. Some participants created their own SS information sheets to give to friends and health care professionals. Creating opportunities to explain difficulties was conducive to receiving valuable social support. Penny's husband had delegated several household tasks to her, which were conflicting with her pacing technique. Penny explained that after discussing the issue with him, he subsequently understood the need to balance activities and that they were able to do this together. Ellie noted the following:

> *I do think that it is helpful to have people that you can talk to about Sjögren's. I mean I have a very close relationship with my family, and I have close friends who I do feel like I can confide in, and that is really helpful for me.* [Ellie]

The freedom and ability to be open and honest about their SS symptoms with trustworthy family members and friends were central to well-being and helped with symptom self-management. However, despite all efforts to communicate effectively, many participants believed that, often, family, friends, and even health care professionals did not fully understand SS. They felt frustrated that symptoms were dismissed, normalized, or incorrectly attributed to other issues such as "getting old" or menopause. Dealing with invisible, ever-changing symptoms was difficult. Multiple general practitioner visits with complaints about seemingly benign symptoms such as fatigue and thirst were sometimes received with skepticism, and the transient nature of these symptoms made the situation worse:

> *Then you're fine and you think, "they'll think I am putting it on."* [Geraldine]

XSL•FO
RenderX

Any respite from symptoms made some participants worry that those around them would not believe them the rest of the time. Carol knew that relative to other conditions that may have 1 visible "major" symptom, her multiple symptoms were unlikely to garner support and understanding because of "Sjögren's [and] all the little things that it has" (Carol). Some participants had stopped attempting to explain their symptoms to family and friends, saying that some symptoms were "very difficult for you to articulate...to somebody who doesn't feel it" (Joan). This was particularly detrimental to relationships with health care professionals. When health care professionals seemed uncompassionate about their symptoms, some participants talked about "shutting down" and making a choice to no longer discuss their SS in consultations. This had negative consequences on participants who ended up feeling rejected and disengaged, and there was a perception that, sometimes, health care practitioners were not even aware of this relational and motivational shift.

When participants felt disbelieved, it led to experiences of self-doubt. Ellie said she was "bounced around like 4-5 practitioners" to the point where she questioned her illness "almost as if it is in your head." Carol resorted to maintaining an activity diary, in part to monitor her fatigue and to preserve her sanity. For her, the diary data provided a sense of external objectivity and an opportunity to feel validated when being questioned by other people:

> *By doing the [diary] you think, yes...I've got a problem and that graph tells me...it is a physical thing, it's not in my mind.* [Carol]

Being diagnosed with SS was a lonely experience for some participants due to the challenges of family, friends, colleagues, and health professionals not relating to participants' symptoms or condition. Social isolation was particularly pronounced for a younger person with SS:

> *I don't know anyone else who has it. So, it is kind of isolating...I also had a hard time finding people who are...my age. So, I mean, I would definitely be interested in meeting younger women who are working, who are finding strategies.* [Ellie]

Overall, connecting with others with SS was important, and participants sought opportunities to meet others with SS, learn, and find the validation and understanding they did not receive from others without SS. Some joined support groups and attended scientific conferences to expand their social circle with other people with SS.

However, not all social contact with others with SS was deemed helpful:

> *Some of the interactions I had honestly more scared me than helped me because it was people who were really in the throes of severe illness and some who weren't coping well, and it was sort of anxiety-provoking.* [Ellie]

Therefore, support from others with SS was generally more welcome when it was helpful and positive, as interactions with those who were struggling to cope could have a negative impact on participants.

Within the construct of relatedness, even positive self-management was found to impact social activity, but having highly supportive friends and family could mitigate this to some extent. Describing and explaining the various, ever-changing symptoms to colleagues, friends, and health professionals who did not fully understand the condition or symptoms could be particularly challenging, but external resources such as using diary data could be a helpful tool to aid communication.

Discussion

Principal Findings

We sought to understand the current self-management approaches used by people with SS to inform the therapeutic ingredients and design recommendations for a self-management smartphone intervention. To date, most studies of lived experience with SS have focused on how specific symptoms are experienced [17,18,48]. To the best of our knowledge, no studies have explored how people with SS perform the day-to-day work of managing their condition and navigating challenges as they do so. This is an important consideration when designing interventions, as those that draw upon users' expertise are more likely to be used [38]. Therefore, we analyzed qualitative data collated through a series of workshops and interviews with people with SS inductively before mapping the themes to the 3 constructs of SDT (competency, autonomy, and relatedness) [52]. This theory was used because it can help identify the psychosocial and practical requirements to support autonomous motivation to adopt and sustain healthy behaviors and to improve well-being in a population [52,67]. Our findings were consistent with what Cartner [68] first described in her qualitative study with participants with SS: the *labor* of living with SS. For her and our participants, competency was an ongoing effort, never a completed achievement. The complex, multisymptomatic, volatile, and unpredictable nature of the condition meant that their hard-earned expertise was being constantly challenged. Having to adapt to an ever-changing and unpredictable challenge evokes the concept of stress, but more specifically, it is captured by the notion of *allostasis:* the work that needs to be performed to find stability within a situation that is constantly changing. When allostasis is frequent or continuous, more work needs to be performed, and our emotional, cognitive, and biological resources can become dysregulated. This is known as *allostatic load*—the psychophysiological wear and tear that occurs to a system that is constantly having to adapt—which has clear links to anxiety, depression, morbidity, and mortality [69].

The labor of the participants and its costs were also evident in the SDT domain of autonomy. Often, there was a degree of forced autonomy, with participants having to perform the epistemic labor of determining how to manage their condition for themselves. This involved ongoing research and even compiling their own resources. Discernment and discrimination were required to determine what advice to trust and follow and how to balance that advice against their own experience. Although this process was enabling, it was also potentially disabling as the process of gathering and compiling information worsened some SS symptoms.

XSL•FO
RenderX

Finally, in the realm of relationships, managing SS requires significant social labor. Often, participants were required to manage the expectations, lack of understanding, skepticism, and disbelief of others, including health professionals, and these efforts were often only partially successful, leading to self-doubt, isolation, and lack of adequate care and support for their illness. This is not dissimilar to the experiences others face with other fatiguing LTCs such as stroke, fibromyalgia, multiple sclerosis, and ankylosing spondylitis [70].

Design Recommendations

Multimedia Appendix 4 [52] summarizes our key findings, which have been mapped to the 3 constructs of SDT, with identified therapeutic approaches and design solutions for each. The findings within these SDT domains were identified as targets for intervention by the participants. In the following sections, we have reviewed these domains and suggested what interventions might help and how the interventions could be incorporated into an app to support self-management.

A key finding within the competency domain was that SS was multisymptomatic, volatile, and unpredictable. Participants were keen for interventions that would impact >1 symptom at a time. A previous study that investigated patient strategies for self-management of inflammatory bowel disease had similar findings [71]. Several treatment approaches and their components discussed during the workshops could potentially address several symptoms simultaneously. For example, activity and sleep management strategies such as pacing and reflective activity diaries have been used to support self-management of pain, sleep disturbances, and fatigue [72-76], and previous studies that evaluated interventions targeting several symptoms have shown promising results. Therefore, we suggest that when designing complex interventions for LTCs, intervention developers should map the potential, identified intervention content to behaviors and symptoms and select techniques that target >1 symptom where possible, thereby placing a smaller demand on the user. While this may not always be possible, streamlining the intervention content where practicable is likely to decrease the possibility of becoming overwhelmed and thereby supports user competency.

The key challenge in the autonomy domain was the amount of work required by participants to determine how to manage their condition on their own. As with many other LTCs, a large part of the "burden of treatment" is shouldered by the person with the condition [77]. Our findings broadly indicate that technology-enabled symptom management could help with this work of illness management. Participants liked the idea of a smartphone app to support self-management. However, merely operationalizing technology is not sufficient to promote and support self-management. Güldenpfennig et al [78] found that poor design and well-meaning paternalism, for example, through automated support that takes active choice away from the user, may compromise autonomy and proactive self-management. Furthermore, intervention designers should aim to strike a more careful balance between the input of experts by experience and those of professional experts [25,78]. In our study, we found that people with SS managed their symptoms using different approaches but that all of them had arrived at their own set of strategies and management regimes through experience, research, and trial and error. Acknowledging the individuality of self-management and the necessity to experiment with different approaches would be a key part of any intervention. Having a repository of strategies in 1 centralized app, which would also allow them to add their own strategies, would seem to be a potentially useful resource. This aligns with previous studies of apps that provide resources while allowing customization and thus may support a user's sense of autonomy [79-82] and move away from a top-down paternalistic or prescriptive approach to LTC management [83,84]. An app for SS would need to combine recognition of the labor of self-management while helping to support it in a manner that honors the user's autonomy and existing wisdom, providing the ability to choose from a range of therapeutic content and to determine the order in which they interact with it.

The most difficult and often fruitless area of labor was observed within the relatedness domain. Participants were required to manage others' expectations, lack of understanding, skepticism, and disbelief, often leading to a smaller social world, isolation, and difficulties in accessing help from health professionals. Again, any intervention needs to begin by acknowledging this labor and the emotional and social costs of having a poorly understood and invisible illness. Our findings also showed that there was often a tension between illness management and maintaining relationships. For example, it could become difficult to implement strategies such as pacing when others were involved, particularly when the person with SS had not fully disclosed their symptoms or condition to the people whom their self-management strategy may affect. Therefore, saying "no" could also be hard for participants, particularly when it was perceived that others would not understand. Other participants had found a solution by working on their means of communicating their difficulties with those around them. Winger et al [85] have found that greater practice of assertiveness and communication skills was associated with reduced pain interference and psychological distress in people with lung cancer, and assertiveness and communication is also a key component of an effective fatigue management intervention for people with rheumatoid arthritis [60]. Therefore, we recommend including assertiveness and communication strategies within a therapeutic self-management app for SS. When considering the design of the app, we recommend including some text to help the user provide a brief explanation about their condition, its symptoms, and their impact to share with health professionals, colleagues, or people in social settings, as needed. We also recommend designing opportunities to practice assertiveness and communication skills within the app for those who may find it helpful.

In summary, our findings suggest that some of the key areas of concern for participants were potentially addressable through an intervention. A common starting point for any approach should be an acknowledgment of the real costs and the daily hard work of having an unpredictable, volatile, and multisymptomatic LTC. Any therapeutic approach needs to be designed to help with this labor; to acknowledge the social, emotional, and physical costs of having and managing SS; and to appreciate the wisdom that the "end user" of the app or

XSL•FO
RenderX

intervention will have already accumulated. Strategies obtained from Acceptance and Commitment Therapy [86] and Compassion Focused Therapy [87] could be useful as they have been used to target the psychosocial impact of other related health complaints such as chronic pain [88]. Next, specific strategies (eg, pacing and sleep management) that could help target multiple symptoms or single symptoms in sequence would be useful. Finally, support to perform some of the social labor involved in living with SS should be a key component. In Multimedia Appendix 4, we have further specified the areas of intervention and suggested the broad therapeutic approaches that might be useful.

Regarding our use of the SDT framework, while it was useful to structure our thinking about intervention development, we also noted that the constructs of SDT often existed in a state of tension with each other, where successfully fulfilling the requirements of one construct leads to reduced functioning in another. As noted previously, this tension occurred between competence and relatedness, where symptom management conflicted with maintaining social bonds. Similar tension existed between autonomy and competence, where participants struggled to feel competent if presented with several self-management options. The SDT states that all 3 fundamental needs have to be met for internally motivated, self-determined behavior to occur [52], but we tentatively suggest that the theory needs to consider moments when some needs stand in opposition to each other. Making the nature of these tensions explicit to the users of an intervention or app would be a key part of its opening narrative.

Limitations

The extent of transferability of our findings to other LTCs is not yet known. However, studies of other autoimmune conditions have demonstrated the same need to self-manage complexity—people with inflammatory bowel disease reported that symptoms (pain, fatigue, and diarrhea) changed over time and could be interconnected at different times, and they required a highly individualized management strategy to “balance” the illness and attend to dynamic fluctuations in symptoms [71]. Overall, our findings may provide insight into how several other autoimmune conditions are self-managed or could be self-managed with the use of an app. However, owing to the nature of the complexity we captured in this study, transferability of our findings to other contexts may be limited. The 17 participants in this study included only 1 (6%) man, which may mean that any unique difficulties experienced by men with SS have been missed in our study. However, SS has a female-to-male incidence rate of 16:1 [89], and the gender makeup of our participants is representative of the wider SS population. Another limitation was that we did not formally collate information about smartphone ownership from participants. Such data should be collated in future similar studies. A final limitation is that most of this study was conducted within the United Kingdom, with only 67% (2/3) of the interviewees living outside the United Kingdom. Therefore, we cannot assume that similar findings would be replicated in other geographical contexts.

Future Studies

Future studies should operationalize the findings of this study to construct an intervention protocol that could be implemented via a smartphone app for the management of SS and empirically optimize its content through pilot and feasibility testing. Furthermore, future studies may explore the transferability of our findings to the self-management contexts of other autoimmune and fluctuating conditions. Our target users were those with primary or secondary SS; future studies should consider how user age influences the design requirements in this patient group.

Conclusions

In conclusion, therapeutic and design approaches for SS should be constructed in both bottom-up (ie, based on the self-management challenges that prospective users already experience) and top-down (according to the most effective treatments documented for SS) formats. For people with SS, choosing to involve an app in their self-management has the possibility of being counterproductive—by adding to their experience of fatigue and becoming overwhelmed. Therefore, the design of a self-management app for SS should support the user in performing the physical, cognitive, emotional, and social work of self-management and should be careful not to add to their already high self-management costs.

Acknowledgments

This study was funded by a grant from Versus Arthritis (grant 22026).

We would like to thank the participants for their time and contribution. In addition we would like to express our gratitude to Michelle Harrison and the North East Sjögren's Syndrome Association, Dr Elizabeth Price, Dr Helen Cartner and Julie Norris for their expert opinion and suggestions regarding the study and the app content.

Conflicts of Interest

None declared.

Multimedia Appendix 1
COREQ (Consolidated Criteria for Reporting Qualitative Research) checklist.
[PDF File (Adobe PDF File), 482 KB - humanfactors_v11i1e54172_app1.pdf]

XSL•FO
RenderX

Multimedia Appendix 2
Workshop topic guide.
[DOCX File , 24 KB - humanfactors_v11i1e54172_app2.docx]

Multimedia Appendix 3
Interview schedule.
[DOCX File , 14 KB - humanfactors_v11i1e54172_app3.docx]

Multimedia Appendix 4
Key findings with potential therapeutic and design recommendations mapped to the Self-Determination Theory domains.
[DOCX File , 23 KB - humanfactors_v11i1e54172_app4.docx]

References

1. Burt J, Roland M, Paddison C, Reeves D, Campbell J, Abel G, et al. Prevalence and benefits of care plans and care planning for people with long-term conditions in England. J Health Serv Res Policy 2012 Jan;17 Suppl 1:64-71. [doi: 10.1258/jhsrp.2011.010172] [Medline: 22315479]
2. Long term conditions compendium of information. Department of Health, United Kingdom Government. 2012 May 30. URL: https://assets.publishing.service.gov.uk/media/5a7c638340f0b62aff6c154e/dh_134486.pdf [accessed 2024-03-28]
3. Care and support for long term conditions. Nuffield Trust. URL: https://tinyurl.com/4h9s46dw [accessed 2023-03-27]
4. Anagnostopoulos I, Zinzaras E, Alexiou I, Papathanasiou AA, Davas E, Koutroumpas A, et al. The prevalence of rheumatic diseases in central Greece: a population survey. BMC Musculoskelet Disord 2010 May 26;11:98 [FREE Full text] [doi: 10.1186/1471-2474-11-98] [Medline: 20504294]
5. Monjardino T, Lucas R, Barros H. Frequency of rheumatic diseases in Portugal: a systematic review. Acta Reumatol Port 2011;36(4):336-363 [FREE Full text] [Medline: 22472925]
6. Chopra A, Abdel-Nasser A. Epidemiology of rheumatic musculoskeletal disorders in the developing world. Best Pract Res Clin Rheumatol 2008 Aug;22(4):583-604. [doi: 10.1016/j.berh.2008.07.001] [Medline: 18783739]
7. Vivino FB. Sjogren's syndrome: clinical aspects. Clin Immunol 2017 Sep;182:48-54. [doi: 10.1016/j.clim.2017.04.005] [Medline: 28428095]
8. Tarn J, Lendrem D, McMeekin P, Lendrem C, Hargreaves B, Ng WF. Primary Sjögren's syndrome: longitudinal real-world, observational data on health-related quality of life. J Intern Med 2022 Jun;291(6):849-855 [FREE Full text] [doi: 10.1111/joim.13451] [Medline: 35018685]
9. Kalouche-Khalil L, Gandhi R, Hao L, Smolkina E, Schumacher F. POS0319 disease burden of primary Sjogren syndrome: a retrospective United States claims database study. Ann Rheum Dis 2022 May 23;81(Suppl 1):410.1-410.1. [doi: 10.1136/annrheumdis-2022-eular.2548]
10. Gairy K, Ruark K, Sinclair SM, Brandwood H, Nelsen L. An innovative online qualitative study to explore the symptom experience of patients with primary Sjögren's syndrome. Rheumatol Ther 2020 Sep;7(3):601-615 [FREE Full text] [doi: 10.1007/s40744-020-00220-9] [Medline: 32725407]
11. Noaiseh G, Baer AN. Toward better outcomes in Sjögren's syndrome: the promise of a stratified medicine approach. Best Pract Res Clin Rheumatol 2020 Feb;34(1):101475. [doi: 10.1016/j.berh.2019.101475] [Medline: 32005417]
12. Brito-Zerón P, Baldini C, Bootsma H, Bowman SJ, Jonsson R, Mariette X, et al. Sjögren syndrome. Nat Rev Dis Primers 2016 Jul 07;2:16047. [doi: 10.1038/nrdp.2016.47] [Medline: 27383445]
13. Miyamoto ST, Lendrem DW, Ng WF, Hackett KL, Valim V. Managing fatigue in patients with primary Sjögren's syndrome: challenges and solutions. Open Access Rheumatol 2019;11:77-88 [FREE Full text] [doi: 10.2147/OARRR.S167990] [Medline: 31118841]
14. Giacomelli C, Talarico R, Baldini C, Bazzichi L. Pain in Sjögren's syndrome. Reumatismo 2014 Jun 06;66(1):39-43 [FREE Full text] [doi: 10.4081/reumatismo.2014.763] [Medline: 24938195]
15. Hackett KL, Gotts ZM, Ellis J, Deary V, Rapley T, Ng WF, et al. An investigation into the prevalence of sleep disturbances in primary Sjögren's syndrome: a systematic review of the literature. Rheumatology (Oxford) 2017 Apr 01;56(4):570-580 [FREE Full text] [doi: 10.1093/rheumatology/kew443] [Medline: 28013207]
16. Cui Y, Xia L, Li L, Zhao Q, Chen S, Gu Z. Anxiety and depression in primary Sjögren's syndrome: a cross-sectional study. BMC Psychiatry 2018 May 16;18(1):131 [FREE Full text] [doi: 10.1186/s12888-018-1715-x] [Medline: 29769121]
17. Hackett KL, Deary V, Deane KH, Newton JL, Ng WF, Rapley T. Experience of sleep disruption in primary Sjögren's syndrome: a focus group study. Br J Occup Ther 2018 Apr;81(4):218-226 [FREE Full text] [doi: 10.1177/0308022617745006] [Medline: 29657352]
18. Ngo DY, Thomson WM, Nolan A, Ferguson S. The lived experience of Sjögren's Syndrome. BMC Oral Health 2016 Feb 01;16:7 [FREE Full text] [doi: 10.1186/s12903-016-0165-4] [Medline: 26831141]

XSL•FO
RenderX

19. Stack RJ, Southworth S, Fisher BA, Barone F, Buckley CD, Rauz S, et al. A qualitative exploration of physical, mental and ocular fatigue in patients with primary Sjögren's syndrome. PLoS One 2017;12(10):e0187272 [FREE Full text] [doi: 10.1371/journal.pone.0187272] [Medline: 29088284]
20. Chen X, Wu H, Wei W. Advances in the diagnosis and treatment of Sjogren's syndrome. Clin Rheumatol 2018 Jul;37(7):1743-1749. [doi: 10.1007/s10067-018-4153-8] [Medline: 29804149]
21. Price EJ, Rauz S, Tappuni AR, Sutcliffe N, Hackett KL, Barone F, et al. The British Society for Rheumatology guideline for the management of adults with primary Sjögren's syndrome. Rheumatology (Oxford) 2017 Oct 01;56(10):1828. [doi: 10.1093/rheumatology/kex375] [Medline: 28957572]
22. Bowman SJ, Everett CC, O'Dwyer JL, Emery P, Pitzalis C, Ng WF, et al. Randomized controlled trial of rituximab and cost-effectiveness analysis in treating fatigue and oral dryness in primary Sjögren's syndrome. Arthritis Rheumatol 2017 Jul;69(7):1440-1450 [FREE Full text] [doi: 10.1002/art.40093] [Medline: 28296257]
23. Gottenberg JE, Ravaud P, Puéchal X, Le Guern V, Sibilia J, Goeb V, et al. Effects of hydroxychloroquine on symptomatic improvement in primary Sjögren syndrome: the JOQUER randomized clinical trial. JAMA 2014 Jul 16;312(3):249-258. [doi: 10.1001/jama.2014.7682] [Medline: 25027140]
24. Hackett KL, Deane KH, Strassheim V, Deary V, Rapley T, Newton JL, et al. A systematic review of non-pharmacological interventions for primary Sjögren's syndrome. Rheumatology (Oxford) 2015 Nov;54(11):2025-2032 [FREE Full text] [doi: 10.1093/rheumatology/kev227] [Medline: 26135587]
25. Hackett KL, Deane KH, Newton JL, Deary V, Bowman SJ, Rapley T, et al. Mixed-methods study identifying key intervention targets to improve participation in daily living activities in primary Sjögren's syndrome patients. Arthritis Care Res (Hoboken) 2018 Jul;70(7):1064-1073 [FREE Full text] [doi: 10.1002/acr.23536] [Medline: 29409110]
26. Barlow J. How to use education as an intervention in osteoarthritis. Best Pract Res Clin Rheumatol 2001 Oct;15(4):545-558. [doi: 10.1053/berh.2001.0172] [Medline: 11567538]
27. Barlow J, Wright C, Sheasby J, Turner A, Hainsworth J. Self-management approaches for people with chronic conditions: a review. Patient Educ Couns 2002;48(2):177-187. [doi: 10.1016/s0738-3991(02)00032-0] [Medline: 12401421]
28. Lorig KR, Holman HR. Self-management education: history, definition, outcomes, and mechanisms. Ann Behav Med 2003 Aug;26(1):1-7. [doi: 10.1207/s15324796abm2601_01]
29. Hackett KL. Developing a non-pharmacological intervention model to improve function and participation in people with primary Sjo¨gren's syndrome. Newcastle University. 2017. URL: https://theses.ncl.ac.uk/jspui/handle/10443/3802 [accessed 2024-03-28]
30. McQueenie R, Jani BD, Siebert S, McLoone P, McCowan C, Macdonald S, et al. Prevalence of chronic pain in LTCs and multimorbidity: a cross-sectional study using UK Biobank. J Multimorb Comorb 2021;11:26335565211005870 [FREE Full text] [doi: 10.1177/26335565211005870] [Medline: 35004337]
31. Morton K, Dennison L, May C, Murray E, Little P, McManus RJ, et al. Using digital interventions for self-management of chronic physical health conditions: a meta-ethnography review of published studies. Patient Educ Couns 2017 Apr;100(4):616-635 [FREE Full text] [doi: 10.1016/j.pec.2016.10.019] [Medline: 28029572]
32. Whitehead L, Seaton P. The effectiveness of self-management mobile phone and tablet apps in long-term condition management: a systematic review. J Med Internet Res 2016 May 16;18(5):e97 [FREE Full text] [doi: 10.2196/jmir.4883] [Medline: 27185295]
33. Wu X, Guo X, Zhang Z. The efficacy of mobile phone apps for lifestyle modification in diabetes: systematic review and meta-analysis. JMIR Mhealth Uhealth 2019 Jan 15;7(1):e12297 [FREE Full text] [doi: 10.2196/12297] [Medline: 30664494]
34. Morrison D, Mair FS, Yardley L, Kirby S, Thomas M. Living with asthma and chronic obstructive airways disease: using technology to support self-management - an overview. Chron Respir Dis 2017 Nov;14(4):407-419 [FREE Full text] [doi: 10.1177/1479972316660977] [Medline: 27512084]
35. Duan H, Wang Z, Ji Y, Ma L, Liu F, Chi M, et al. Using goal-directed design to create a mobile health app to improve patient compliance with hypertension self-management: development and deployment. JMIR Mhealth Uhealth 2020 Feb 25;8(2):e14466 [FREE Full text] [doi: 10.2196/14466] [Medline: 32130161]
36. Wulfovich S, Fiordelli M, Rivas H, Concepcion W, Wac K. "I must try harder": design implications for mobile apps and wearables contributing to self-efficacy of patients with chronic conditions. Front Psychol 2019 Oct 23;10:2388 [FREE Full text] [doi: 10.3389/fpsyg.2019.02388] [Medline: 31749733]
37. Osborne CL, Juengst SB, Smith EE. Identifying user-centered content, design, and features for mobile health apps to support long-term assessment, behavioral intervention, and transitions of care in neurological rehabilitation: An exploratory study. Brit J Occup Ther 2020 Oct 07;84(2):101-110. [doi: 10.1177/0308022620954115]
38. Druce KL, Dixon WG, McBeth J. Maximizing engagement in mobile health studies: lessons learned and future directions. Rheum Dis Clin North Am 2019 May;45(2):159-172 [FREE Full text] [doi: 10.1016/j.rdc.2019.01.004] [Medline: 30952390]
39. Maharaj A, Lim D, Murphy R, Serlachius A. Comparing two commercially available diabetes apps to explore challenges in user engagement: randomized controlled feasibility study. JMIR Form Res 2021 Jun 16;5(6):e25151 [FREE Full text] [doi: 10.2196/25151] [Medline: 34132640]

XSL•FO
RenderX

40. Yardley L, Spring BJ, Riper H, Morrison LG, Crane DH, Curtis K, et al. Understanding and promoting effective engagement with digital behavior change interventions. Am J Prev Med 2016 Nov;51(5):833-842. [doi: 10.1016/j.amepre.2016.06.015] [Medline: 27745683]
41. Morrison LG. Theory-based strategies for enhancing the impact and usage of digital health behaviour change interventions: a review. Digit Health 2015;1:2055207615595335 [FREE Full text] [doi: 10.1177/2055207615595335] [Medline: 29942544]
42. Peters D, Davis S, Calvo RA, Sawyer SM, Smith L, Foster JM. Young people's preferences for an asthma self-management app highlight psychological needs: a participatory study. J Med Internet Res 2017 Apr 11;19(4):e113 [FREE Full text] [doi: 10.2196/jmir.6994] [Medline: 28400353]
43. Yardley L, Morrison L, Bradbury K, Muller I. The person-based approach to intervention development: application to digital health-related behavior change interventions. J Med Internet Res 2015 Jan 30;17(1):e30 [FREE Full text] [doi: 10.2196/jmir.4055] [Medline: 25639757]
44. Rogers Y, Marshall P. Research in the wild. In: Synthesis Lectures on Human-Centered Informatics. San Rafael, CA: Morgan & Claypool Publishers; Apr 04, 2017:i-97.
45. Morris ME. Preventing falls in older people. BMJ 2012 Aug 07;345:e4919. [doi: 10.1136/bmj.e4919] [Medline: 22872696]
46. van Middelaar T, Beishuizen CR, Guillemont J, Barbera M, Richard E, Moll van Charante EP. Engaging older people in an internet platform for cardiovascular risk self-management: a qualitative study among Dutch HATICE participants. BMJ Open 2018 Jan 21;8(1):e019683 [FREE Full text] [doi: 10.1136/bmjopen-2017-019683] [Medline: 29358447]
47. Schermer M. Telecare and self-management: opportunity to change the paradigm? J Med Ethics 2009 Nov;35(11):688-691. [doi: 10.1136/jme.2009.030973] [Medline: 19880706]
48. Lackner A, Ficjan A, Stradner MH, Hermann J, Unger J, Stamm T, et al. It's more than dryness and fatigue: the patient perspective on health-related quality of life in primary Sjögren's Syndrome - a qualitative study. PLoS One 2017;12(2):e0172056 [FREE Full text] [doi: 10.1371/journal.pone.0172056] [Medline: 28182787]
49. Sanz MF, Acha BV, García MF. Co-design for people-centred care digital solutions: a literature review. Int J Integr Care 2021 Apr 30;21(2):16 [FREE Full text] [doi: 10.5334/ijic.5573] [Medline: 33981193]
50. Grenier C, Ibrahim R, Paixão-Barradas S. Designing and innovating in digital healthcare: co-design for taking patients' needs into account. In: Béranger J, Rizoulières R, editors. The Digital Revolution in Health. Hoboken, NJ: John Wiley & Sons; 2021.
51. Harrington CN, Wilcox L, Connelly K, Rogers W, Sanford J. Designing health and fitness apps with older adults: examining the value of experience-based co-design. In: Proceedings of the 12th EAI International Conference on Pervasive Computing Technologies for Healthcare. 2018 Presented at: PervasiveHealth '18; May 21-24, 2018; New York, NY. [doi: 10.1145/3240925.3240929]
52. Ryan RM, Deci EL. Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. Am Psychol 2000 Jan;55(1):68-78. [doi: 10.1037//0003-066x.55.1.68] [Medline: 11392867]
53. Gillison FB, Rouse P, Standage M, Sebire SJ, Ryan RM. A meta-analysis of techniques to promote motivation for health behaviour change from a self-determination theory perspective. Health Psychol Rev 2019 Mar;13(1):110-130. [doi: 10.1080/17437199.2018.1534071] [Medline: 30295176]
54. Peters D, Calvo RA, Ryan RM. Designing for motivation, engagement and wellbeing in digital experience. Front Psychol 2018;9:797 [FREE Full text] [doi: 10.3389/fpsyg.2018.00797] [Medline: 29892246]
55. Eassey D, Reddel HK, Ryan K, Smith L. 'It is like learning how to live all over again' a systematic review of people's experiences of living with a chronic illness from a self-determination theory perspective. Health Psychol Behav Med 2020 Jul 23;8(1):270-291 [FREE Full text] [doi: 10.1080/21642850.2020.1794879] [Medline: 34040872]
56. Stenberg N, Gillison F, Rodham K. How do peer support interventions for the self-management of chronic pain, support basic psychological needs? a systematic review and framework synthesis using self-determination theory. Patient Educ Couns 2022 Nov;105(11):3225-3234 [FREE Full text] [doi: 10.1016/j.pec.2022.07.017] [Medline: 35985906]
57. Hekler EB, Klasnja P, Riley WT, Buman MP, Huberty J, Rivera DE, et al. Agile science: creating useful products for behavior change in the real world. Transl Behav Med 2016 Jun;6(2):317-328 [FREE Full text] [doi: 10.1007/s13142-016-0395-7] [Medline: 27357001]
58. Tong A, Sainsbury P, Craig J. Consolidated Criteria for Reporting Qualitative Research (COREQ): a 32-item checklist for interviews and focus groups. Int J Qual Health Care 2007 Dec;19(6):349-357. [doi: 10.1093/intqhc/mzm042] [Medline: 17872937]
59. Andersen K, Wakkary R. The magic machine workshops: making personal design knowledge. In: Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems. 2019 Presented at: CHI '19; May 4-9, 2019; Glasgow, UK. [doi: 10.1145/3290605.3300342]
60. Hewlett S, Almeida C, Ambler N, Blair PS, Choy EH, Dures E, RAFT Study Group. Reducing arthritis fatigue impact: two-year randomised controlled trial of cognitive behavioural approaches by rheumatology teams (RAFT). Ann Rheum Dis 2019 Apr;78(4):465-472 [FREE Full text] [doi: 10.1136/annrheumdis-2018-214469] [Medline: 30793700]
61. Hackett KL, Lambson RL, Strassheim V, Gotts Z, Deary V, Newton JL. A concept mapping study evaluating the UK's first NHS generic fatigue clinic. Health Expect 2016 Oct;19(5):1138-1149 [FREE Full text] [doi: 10.1111/hex.12405] [Medline: 26332418]

XSL•FO
RenderX

62. Ericsson KA, Simon HA. Verbal reports as data. Psychol Rev 1980 May;87(3):215-251. [doi: 10.1037/0033-295X.87.3.215]
63. Proudfoot K. Inductive/deductive hybrid thematic analysis in mixed methods research. J Mixed Method Res 2022 Sep 20;17(3):308-326. [doi: 10.1177/15586898221126816]
64. Braun V, Clarke V. Successful Qualitative Research: A Practical Guide for Beginners. Thousand Oaks, CA: SAGE Publications; Mar 22, 2013.
65. Deci EL, Ryan RM. Self-determination theory: a macrotheory of human motivation, development, and health. Can Psychol 2008 Aug;49(3):182-185. [doi: 10.1037/a0012801]
66. Ryan RM, Deci EL. Chapter four - brick by brick: the origins, development, and future of self-determination theory. In: Elliot AJ, editor. Advances in Motivation Science. Cambridge, MA: Academic Press; 2019.
67. Ntoumanis N, Ng JY, Prestwich A, Quested E, Hancox JE, Thøgersen-Ntoumani C, et al. A meta-analysis of self-determination theory-informed intervention studies in the health domain: effects on motivation, health behavior, physical, and psychological health. Health Psychol Rev 2021 Jun;15(2):214-244 [FREE Full text] [doi: 10.1080/17437199.2020.1718529] [Medline: 31983293]
68. Cartner H. 'A Taste of my Life'. Altered Eating and Sjogren's Syndrome: A Participant-led Exploration of Eating Experience Using Application of a Novel Methodology, 'Food Play', in Department of Psychology 2024;Northumbria University:1-371 (forthcoming).
69. Parker HW, Abreu AM, Sullivan MC, Vadiveloo MK. Allostatic load and mortality: a systematic review and meta-analysis. Am J Prev Med 2022 Jul;63(1):131-140. [doi: 10.1016/j.amepre.2022.02.003] [Medline: 35393143]
70. Eilertsen G, Ormstad H, Kirkevold M, Mengshoel AM, Söderberg S, Olsson M. Similarities and differences in the experience of fatigue among people living with fibromyalgia, multiple sclerosis, ankylosing spondylitis and stroke. J Clin Nurs 2015 Jul;24(13-14):2023-2034. [doi: 10.1111/jocn.12774] [Medline: 25661994]
71. Dibley L, Khoshaba B, Artom M, Van Loo V, Sweeney L, Syred J, et al. Patient strategies for managing the vicious cycle of fatigue, pain and urgency in inflammatory bowel disease: impact, planning and support. Dig Dis Sci 2021 Oct;66(10):3330-3342. [doi: 10.1007/s10620-020-06698-1] [Medline: 33164146]
72. Andrews NE, Strong J, Meredith PJ, Branjerdporn GS. Approach to activity engagement and differences in activity participation in chronic pain: a five-day observational study. Aust Occup Ther J 2018 Dec;65(6):575-585. [doi: 10.1111/1440-1630.12516] [Medline: 30178871]
73. Carney CE, Buysse DJ, Ancoli-Israel S, Edinger JD, Krystal AD, Lichstein KL, et al. The consensus sleep diary: standardizing prospective sleep self-monitoring. Sleep 2012 Feb 01;35(2):287-302 [FREE Full text] [doi: 10.5665/sleep.1642] [Medline: 22294820]
74. Thomas S, Thomas PW, Nock A, Slingsby V, Galvin K, Baker R, et al. Development and preliminary evaluation of a cognitive behavioural approach to fatigue management in people with multiple sclerosis. Patient Educ Couns 2010 Feb;78(2):240-249. [doi: 10.1016/j.pec.2009.07.001] [Medline: 19665337]
75. Antcliff D, Keeley P, Campbell M, Oldham J, Woby S. The development of an activity pacing questionnaire for chronic pain and/or fatigue: a Delphi technique. Physiotherapy 2013 Sep;99(3):241-246. [doi: 10.1016/j.physio.2012.12.003] [Medline: 23395263]
76. McCrae CS, Williams J, Roditi D, Anderson R, Mundt JM, Miller MB, et al. Cognitive behavioral treatments for insomnia and pain in adults with comorbid chronic insomnia and fibromyalgia: clinical outcomes from the SPIN randomized controlled trial. Sleep 2019 Mar 01;42(3):1-15 [FREE Full text] [doi: 10.1093/sleep/zsy234] [Medline: 30496533]
77. Mair FS, May CR. Thinking about the burden of treatment. BMJ 2014 Nov 10;349:g6680. [doi: 10.1136/bmj.g6680] [Medline: 25385748]
78. Güldenpfennig F, Mayer P, Panek P, Fitzpatrick G. An autonomy-perspective on the design of assistive technology experiences of people with multiple sclerosis. In: Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems. 2019 Presented at: CHI '19; May 4-9, 2019; Glasgow, UK. [doi: 10.1145/3290605.3300357]
79. Seah ML, Koh KT. The efficacy of using mobile applications in changing adolescent girls' physical activity behaviour during weekends. Eur Phys Educ Rev 2020 Jun 17;27(1):113-131. [doi: 10.1177/1356336x20930741]
80. Marathe S, Shyam Sundar S. What drives customization?: control or identity? In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. 2011 Presented at: CHI '11; May 7-12, 2011; Vancouver, BC. [doi: 10.1145/1978942.1979056]
81. Harrison D, Marshall P, Bianchi-Berthouze N, Bird J. Activity tracking: barriers, workarounds and customisation. In: Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing. 2015 Presented at: UbiComp '15; September 7-11, 2015; Osaka, Japan. [doi: 10.1145/2750858.2805832]
82. Zhang R, Ringland KE, Paan M, Mohr DC, Reddy M. Designing for emotional well-being: integrating persuasion and customization into mental health technologies. In: Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems. 2021 Presented at: CHI '21; May 8-13, 2021; Yokohama, Japan. [doi: 10.1145/3411764.3445771]
83. Carswell C, Brown JV, Lister J, Ajjan RA, Alderson SL, Balogun-Katung A, et al. The lived experience of severe mental illness and long-term conditions: a qualitative exploration of service user, carer, and healthcare professional perspectives on self-managing co-existing mental and physical conditions. BMC Psychiatry 2022 Jul 19;22(1):479 [FREE Full text] [doi: 10.1186/s12888-022-04117-5] [Medline: 35850709]

XSL•FO
RenderX

84. Howell D, Mayer DK, Fielding R, Eicher M, Verdonck-de Leeuw IM, Johansen C, et al. Management of cancer and health after the clinic visit: a call to action for self-management in cancer care. J Natl Cancer Inst 2021 May 04;113(5):523-531 [FREE Full text] [doi: 10.1093/jnci/djaa083] [Medline: 32525530]
85. Winger JG, Rand KL, Hanna N, Jalal SI, Einhorn LH, Birdas TJ, et al. Coping skills practice and symptom change: a secondary analysis of a pilot telephone symptom management intervention for lung cancer patients and their family caregivers. J Pain Sympt Manag 2018 May;55(5):1341-9.e4 [FREE Full text] [doi: 10.1016/j.jpainsymman.2018.01.005] [Medline: 29366911]
86. Hayes SC, Lillis J. Acceptance and Commitment Therapy. Washington, DC: American Psychological Association; 2012.
87. Gilbert P. Introducing compassion-focused therapy. Adv Psychiatr Treat 2018 Jan 02;15(3):199-208. [doi: 10.1192/apt.bp.107.005264]
88. McCracken LM, Yu L, Vowles KE. New generation psychological treatments in chronic pain. BMJ 2022 Feb 28;376:e057212. [doi: 10.1136/bmj-2021-057212] [Medline: 35228207]
89. Brandt JE, Priori R, Valesini G, Fairweather D. Sex differences in Sjögren's syndrome: a comprehensive review of immune mechanisms. Biol Sex Differ 2015;6:19 [FREE Full text] [doi: 10.1186/s13293-015-0037-7] [Medline: 26535108]

Abbreviations

COREQ: Consolidated Criteria for Reporting Qualitative Research
LTC: long-term condition
SDT: Self-Determination Theory
SS: Sjögren's syndrome

Edited by C Jacob; submitted 01.11.23; peer-reviewed by S Stones, A Chen, R Grainger; comments to author 26.12.23; revised version received 26.01.24; accepted 31.01.24; published 17.04.24.

Please cite as:
McCallum C, Campbell M, Vines J, Rapley T, Ellis J, Deary V, Hackett K
A Smartphone App to Support Self-Management for People Living With Sjögren's Syndrome: Qualitative Co-Design Workshops
JMIR Hum Factors 2024;11:e54172
URL: https://humanfactors.jmir.org/2024/1/e54172
doi:10.2196/54172
PMID:38630530

©Claire McCallum, Miglena Campbell, John Vines, Tim Rapley, Jason Ellis, Vincent Deary, Katie Hackett. Originally published in JMIR Human Factors (https://humanfactors.jmir.org), 17.04.2024. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on https://humanfactors.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Effective Communication Supported by an App for Pregnant Women: Quantitative Longitudinal Study

Lukas Kötting[1], MSc; Vinayak Anand-Kumar[1], MSc; Franziska Maria Keller[2], MSc, PhD; Nils Tobias Henschel[1], MSc; Sonia Lippke[1], Dr

[1]Psychology and Methods, School of Business, Social & Decision Sciences, Constructor University Bremen gGmbH, Bremen, Germany
[2]Klinikum Bremerhaven Reinkenheide gGmbH, Bremerhaven, Germany

Corresponding Author:
Sonia Lippke, Dr
Psychology and Methods
School of Business, Social & Decision Sciences
Constructor University Bremen gGmbH
Campus Ring 1
Bremen, 28759
Germany
Phone: 49 421 200 4730
Email: s.lippke@jacobs-university.de

Abstract

Background: In the medical field of obstetrics, communication plays a crucial role, and pregnant women, in particular, can benefit from interventions improving their self-reported communication behavior. Effective communication behavior can be understood as the correct transmission of information without misunderstanding, confusion, or losses. Although effective communication can be trained by patient education, there is limited research testing this systematically with an app-based digital intervention. Thus, little is known about the success of such a digital intervention in the form of a web-app, potential behavioral barriers for engagement, as well as the processes by which such a web-app might improve self-reported communication behavior.

Objective: This study fills this research gap by applying a web-app aiming at improving pregnant women's communication behavior in clinical care. The goals of this study were to (1) uncover the potential risk factors for early dropout from the web-app and (2) investigate the social-cognitive factors that predict self-reported communication behavior after having used the web-app.

Methods: In this study, 1187 pregnant women were recruited. They all started to use a theory-based web-app focusing on intention, planning, self-efficacy, and outcome expectancy to improve communication behavior. Mechanisms of behavior change as a result of exposure to the web-app were explored using stepwise regression and path analysis. Moreover, determinants of dropout were tested using logistic regression.

Results: We found that dropout was associated with younger age (P=.014). Mechanisms of behavior change were consistent with the predictions of the health action process approach. The stepwise regression analysis revealed that action planning was the best predictor for successful behavioral change over the course of the app-based digital intervention (β=.331; P<.001). The path analyses proved that self-efficacy beliefs affected the intention to communicate effectively, which in turn, elicited action planning and thereby improved communication behavior (β=.017; comparative fit index=0.994; Tucker–Lewis index=0.971; root mean square error of approximation=0.055).

Conclusions: Our findings can guide the development and improvement of apps addressing communication behavior in the following ways in obstetric care. First, such tools would enable action planning to improve communication behavior, as action planning is the key predictor of behavior change. Second, younger women need more attention to keep them from dropping out. However, future research should build upon the gained insights by conducting similar internet interventions in related fields of clinical care. The focus should be on processes of behavior change and strategies to minimize dropout rates, as well as replicating the findings with patient safety measures.

Trial Registration: ClinicalTrials.gov identifier: NCT03855735; https://classic.clinicaltrials.gov/ct2/show/NCT03855735

(JMIR Hum Factors 2024;11:e48218) doi:10.2196/48218

XSL•FO
RenderX

KEYWORDS

clinical care; health action process approach; HAPA; intention; communication behavior; patient safety; patient education; internet intervention; dropout; digital health; behavior change; prediction; obstetric; pregnant women; pregnancy; safe communication; health behaviors; obstetric care

Introduction

Background

In the dynamic landscape of medical internet research, the pursuit of effective interventions and preventive health programs demands a comprehensive understanding of diverse populations, including pregnant women and their unique needs. This paper unveils the outcomes of formative research and preliminary results within the realm of medical and preventive health, exploring an innovation and technology in terms of a digital intervention, that is, apps aiming at improving communication behavior. Formative research, characterized by its emphasis on gathering insights from the intended beneficiaries, emerges as a fundamental tool for tailoring interventions to meet the unique requirements of diverse communities [1]. By demonstrating the integral role of formative research in the early stages of program development, we aim to provide a compelling case for its incorporation in the toolkit of researchers and experts working in the field of medical internet research. In this paper, we outline the potential of using digital tools like apps for improving communication behavior for patient safety and the risks involved with regard to dropouts in app-based interventions. Lastly, we outline a behavior change theory to model communication behavior, which may help to map out the health-related behavior more systematically and which was used for our research investigating communication behavior and app usage.

Patient Safety in Health Care

Patient safety is defined as the absence of harm that could have been prevented in patients. For achieving patient safety, health care should be delivered in an optimal manner, trust should be built among all involved individuals, and misunderstanding, information loss, or error occurrence should be prevented [2]. Thus, patient safety requires effective communication behaviors among health care professionals, patients, their partners, or accompanying persons [2]. In particular, in obstetric care, this holds true [2,3] because women in labor have to express their needs and wishes even in the face of stress and barriers to ensure their active role in the obstetric process. Communication behavior can be measured and taught [3-5] and is a reliable approach for improving patient safety [6,7]. Communication behavior involves multiple individuals, including patients, health care workers, and partners [8-13]. This encompasses not only the importance of perceiving a supportive environment that guarantees an open exchange of concerns and potential solutions but also the individual's competency to communicate safely. Such competency consists of the self-reported communication skills that are based on Rider and Keefer's [14] competencies and are impacted by determinants of the communication behavior, that is, self-efficacy, intention formation, and planning [15-19].

Communication Behavior

Communication is defined as a process involving the exchange of cognitions and emotions through verbal and nonverbal actions [20,21]. In this work, we define patients' communication similarly to communication in health care workers to keep definitions for both groups aligned. Previous work [8,10] performed over the scope of this project has defined communication in line with Rider and Keefer's definition [14]. They describe a set of skills including the creation and sustainability of a therapeutic relationship, use of effective listening, prompt and effective responding, and effective communication [14]. Effective communication is the correct transmission of information without misunderstanding, confusion, or losses.

Although effective communication has been shown to be of importance in preventing errors in medical care as well as in patient-provider relationships [6,22-24], only few studies have investigated effective communication behavior among those receiving obstetric care. Moreover, there is limited evidence for innovative tools aiming at increasing effective communication among pregnant women and their support networks [8,11,25]. Previous research has mainly investigated face-to-face interventions in clinical care or hospital settings [26,27]. Although traditional face-to-face interventions demonstrate efficacy, they tend to show several disadvantages concerning feasibility, such as higher financial constraints, limited utilization due to mobility constraints, or scheduling and time issues [28-31]. These constraints of traditional face-to-face interventions also call for cost-effective, convenient, instantly available, and scalable alternative solutions. One of these alternatives, successfully implemented across multiple therapeutic areas, including the promotion of health behavior change, is support via the internet, digital interventions, and apps in the medical field [32-35].

Digital Interventions and Apps

Digital interventions and apps (also called as medical internet support or web-based communication training) have shown several advantages over traditional face-to-face interventions, such as increased ease of accessibility and personalized interactions with real-time feedback. Furthermore, they offer the opportunity for scalability to larger populations, including individuals who live in remote areas. Moreover, such digital interventions can be relatively cost-effective compared to traditional formats [36,37]. Although there is clear general evidence regarding digital interventions, there is scarcity of research on those targeting to foster effective communication.

The same holds true regarding the applicability and integrability of traditional health behavior change theories such as the health action process approach (HAPA) to explain health behavior changes in digital interventions (ie, smartphone apps). Indeed, literature shows how interventions supporting motivational and volitional processes prove effective [8,9,38]. However, the

XSL•FO
RenderX

HAPA model has been rarely applied to interventions targeting effective communication [8,9], and it is hardly ever used to explain communication behavior in the context of digital interventions or their dropout of pregnant women. Therefore, we review dropout in more detail.

Factors Associated With App Usage and Dropout From Digital Interventions

Early dropout from digital interventions is a key problem [39], as the intervention use is discontinued. This needs more attention because if users drop out, which might occur as often as 1 in 2 cases [40], efficacy is limited, and the reach and generalizability of the obtained results are diminished [39]. However, little is known about the factors associated with dropout [39]. Accordingly, more research investigating and identifying such factors is needed, especially in the context of communication behavior and giving birth.

Looking at the general literature on the potential risk factors for early dropout in digital interventions, the following sociodemographic and behavior change factors were identified: age, education, and social support [41,42]. Although there is no previous study on dropout from digital interventions addressing effective communication of pregnant women, evidence from other areas with digital interventions exist. For example, Wu and colleagues [43] investigated dropout in a blended care cognitive behavioral intervention. They highlighted that a higher dropout rate was associated specifically with female gender, poorer financial status, and the absence of a college degree. Additionally, Gao et al [44] found that younger patients and those who were less educated were more likely to drop out from digital intervention studies. Other factors associated with early dropout were marital status (higher probability of divorced individuals to drop out) and ethnicity [45]. Besides the sociodemographic factors, according to Davis and Addis [46], psychological determinants should also be considered while examining dropouts from digital health interventions. According to [47,48], users with low intention to change their behavior have been found to drop out more often from digital interventions. A study by Schroé and colleagues [49] further investigated why users discontinued the use of digital health interventions. Their results highlighted that whereas sociodemographic factors were predictive of early dropout, psychological determinants such as action planning and self-monitoring were associated with completion of digital interventions [49]. This is in line with other research highlighting that self-monitoring [50] and higher intrinsic motivation were associated with lower attrition rates [51]. A theory that could bring the different factors together to enable systematic research is the aforementioned HAPA model, which is described in more detail below.

HAPA Model to Understand and Improve Behavior Change

Self-reported communication behavior constitutes a preventive health behavior [25] and may be fostered by the same factors and processes that health psychology literature has repeatedly showcased [52-54]. HAPA proved to be a useful theory [11] essentially since it considers the interplay of resources, barriers, as well as the well-known behavior intention gap [54,55]. The HAPA model is divided into 2 distinct phases: (1) the motivational phase in which individuals consider their competencies' determinants such as self-efficacy, expectations about behavioral outcomes (outcome expectancies) and formulate a behavioral intention (eg, to communicate in the birthing context), and (2) the volitional phase, wherein pregnant women develop and enact behavioral plans in order to bring the intentions to behavioral actions. This whole process is shaped by social-cognitive barriers and facilitators that may originate externally or stem from women's personal belief, which is also called self-efficacy [54,56]. According to the HAPA model, individuals need to first form an intention, which is based on outcome expectancies and self-efficacy, before acting accordingly. Hence, the pathway of intention on the actual behavior is mediated by action planning [8,11,57]—with action planning being more proximal to behavior, and intention, outcome expectancies, and self-efficacy being more distal to behavior.

In order to improve communication behavior, interventions must be tailored to social-cognitive barriers and facilitators of the target population. Previous evidence has demonstrated that classical face-to-face interventions based on motivational and volitional theories such as HAPA are effective in improving self-reported communication [8,9,38]. It should be noted that most of these findings stem from interventions that were solely offered to health care workers [25], but more attention needs to be paid to patient education. This is the basis of our study with pregnant women randomized into an intervention group or a waitlist control group.

Goal of This Study

As previously outlined, there is a need for further studies to investigate effective communication behaviors of pregnant women within the context of a digital intervention. The goals of our study were 2-fold. First, we aimed to uncover the potential risk factors for early dropout from a digital intervention. Second, we aimed to investigate the social-cognitive factors that would predict the self-reported communication behavior after having used the digital intervention. Thus, the hypotheses are as follows:

1. Hypothesis 1: Sociodemographic factors play a larger role in predicting dropout during a digital intervention relative to behavior change variables.
2. Hypothesis 2: The social-cognitive factors outlined by HAPA (self-efficacy, outcome expectancy, and action planning) predict self-reported communication behavior in pregnant women over the course of the app-based intervention.
3. Hypothesis 3: More distal HAPA variables (intention, outcome expectancies, and self-efficacy) indirectly relate with self-reported communication behavior mediated by action planning.

Methods

TeamBaby Project

This study stems from a larger project named TeamBaby, which was tasked with developing interventions to improve

XSL•FO
RenderX

communication behavior between those who receive and provide obstetric care. One of the interventions was a digital intervention, that is, an app (actually, a web-app). Data collected from the TeamBaby web-app were used to investigate our hypotheses. The TeamBaby project was funded by the German Innovation Fund (project 01VSF18023) of the Gemeinsamer Bundesausschuss and preregistered (ClinicalTrials.gov identifier: NCT03855735) on February 27, 2019.

Recruitment and Procedures

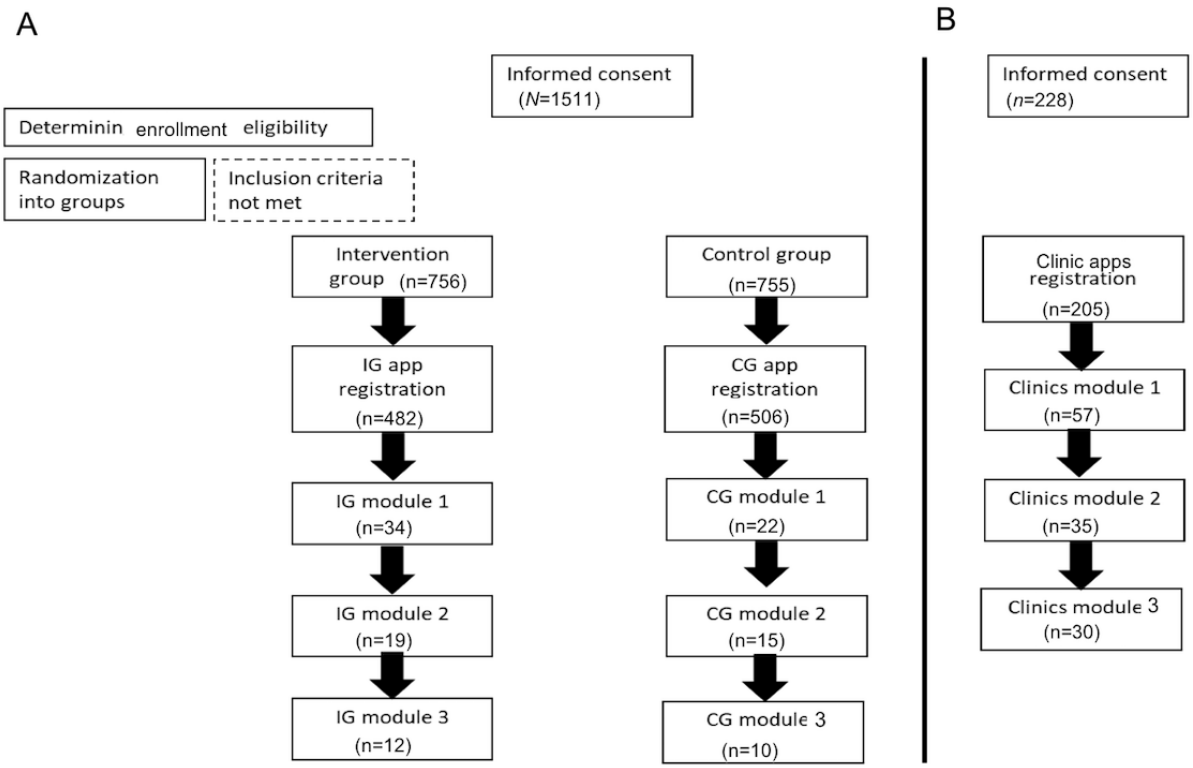
Ethics Approval

Ethics approval for data collection in the maternity clinics was granted by the ethics committee for human research of the University Hospital Ulm (114/19) and the ethics committee for medical research of the University Hospital Frankfurt am Main (19-292). Informed consent was provided in the registration process, and all data were anonymized by providing users with a random ID that could not be linked to user emails or personal IDs. No compensation was provided for participation in this study.

Participants

Participants were recruited into this study, as outlined in Figure 1. Participants represented a pragmatic sample. Sample calculations were performed prior to data collection for an assumed dropout of 20%. We estimated that 176 or more individuals would be needed to recruit [9]. All recruited women were able to register to use the TeamBaby web-app if they were residing in Germany, either during the time of our study (if randomized into the intervention group) or 2 weeks later (if randomized into the waitlist-control arm), that is, pregnant women and their support persons.

Figure 1. Flowchart of the study participants. (A) Study participants using the app as intervention group versus being randomized to the control group and using the app only 2 weeks later. (B) Study flow for the clinic's intervention group. CG: control group; IG: intervention group.



Pregnant women were recruited through 1 of the 2 recruitment channels (Figure 1): (1) those who sought treatment during pregnancy in project-affiliated clinics or (2) those who were currently pregnant and based anywhere in Germany. With respect to the former group, project members worked together with obstetricians to recruit patients. The Germany-wide recruitment utilized social media and targeted advertising to promote this study. In addition, flyers were placed in health care clinics and pharmacies across the country. Women were eligible to participate if they were currently pregnant, had sufficient knowledge of German, and were at least 18 years old.

Participants were recruited into this study upon completing a web-based questionnaire. Women recruited in the clinics were invited to register with the web-app. Women from the Germany-wide recruitment were randomized into either a treatment group or waitlist-control group. The treatment group was presented with a link to use the web-app directly. The waitlist-control arm was provided with a link to the web-app 2 weeks later. All users of the web-app were presented with a series of questions at regular timepoints to determine whether there were changes in HAPA variables and communication behavior as users progressed through the 3 modules of the web-app. Participants who completed less than 2 modules were considered as early dropouts.

Intervention Content

The TeamBaby web-app provided guidance on how to work effectively with health care workers. The web-app consisted of 10 lessons, wrapped in 3 modules, which were developed and

XSL•FO
RenderX

structured based on the processes of behavior change as set out by the HAPA model. The first set of lessons was designed to increase the outcome expectations of effective communication behavior and create an intention to adopt self-reported communication practices. The subsequent lessons were designed to increase the belief or trust in oneself to employ effective communication and enable users to make tangible plans for implementing self-reported communication behavior. More detailed information about the modules and the content can be found in Multimedia Appendix 1.

Measures

Participants were asked to complete questions relating to communication behavior and HAPA variables at 4 timepoints: before starting the first module and after completing each subsequent module. The assessment of self-reported communication behavior was based on Rider and Keefer's competencies [14] and adapted to address pregnant women's behavior in previous publications [12]. As the aim of the research was to understand the underlying social-cognitive processes of communication, the items were developed to capture self-reported communication behavior. Table 1 presents the items [14,58,59] used to evaluate each variable.

Table 1. Measured health action process approach and self-reported communication variables.

Variables	Item example	Range[a]
Communication behavior (7 items) [14]	During pregnancy, I always have communicated my needs clearly.	1-6
Outcome expectancy (single item) [58]	If I communicate well with doctors and midwives, my preferences can be considered during childbirth.	1-6
Intention (single item) [59]	I intend to always make sure that I communicate effectively with the doctors and midwives.	1-6
Self-efficacy (single item) [59]	I am sure I can communicate well even when I am tired or exhausted.	1-6
Action planning (single item) [58]	I have planned precisely how to communicate well while giving birth.	1-6

[a]Range of 1-6 spans from "does not apply at all" to "does apply fully."

Sociodemographic Data

In addition to behavior change measures, demographic information was collected. Participants reported their age, marital status, highest level of education, and nationality.

Aggregated Variables

Participant communication scores were combined into a single item for each individual by taking the average across the different communication behaviors. This was expressed as overall communication. The implicit assumption here is that more effective communication within the described obstetric setting should facilitate a safer birth.

Statistical Analysis

All analyses were conducted using the R [60] and RStudio [61] software. Significance was determined at the 5% level. The aim was to determine what variables would predict early dropout from the web-app. Early dropout was expressed as a binary variable: participants were marked as dropping out early if less than 2 modules were completed. For example, participants who completed 2 or 3 modules were marked as 0 (ie, not dropped out early), while participants who completed only 1 module or none were marked as 1 (ie, dropped out early).

To investigate hypothesis 1, a general logistic regression model using the glm function was built to identify whether HAPA variables, age, marital status, education, and recruitment channel predicted dropout. Recruitment channel was a categorical variable that reflected entry into the app. Participants entered the app either directly through clinical recruitment or in the Germany-wide recruitment after randomization into the intervention group or after the waiting time when being randomized into the waitlist control arm. In the logistic regression model, the clinical recruitment group was the reference group to which the other recruitment channels were compared to.

A hierarchical regression was performed to investigate hypothesis 2, using the "lm" function available in Base R; HAPA variables were sequentially added to build a final model that predicts post web-app communication. Table 2 outlines how the predictor and outcome variables were operationalized in the model. Following the construction of a model to explain changes in post web-app communication, possible processes for behavior change were proposed. To investigate hypothesis 3, a structural equation model using the lavaan package [62] was built to identify how HAPA variables related to one another and in turn contributed to changes in post web-app communication.

XSL•FO
RenderX

Table 2. Overview of the operationalization of social-cognitive predictors and communication behavior.

Type	Variable	Operationalization	Example
Outcome	• Post web-app communication, expressed as C_t	Respondent's most recent overall communication score after having completed at least 1 module of the web-app. Only individuals who completed at least 1 module were included in the analyses to ensure the data captured those that used the web-app.	If an individual had communication scores after modules 1, 2, and 3, only the response after module 3 was used.
Predictor	• Outcome expectancy at the preceding timepoint, expressed as OE_{t-1} • Intention at the preceding timepoint, expressed as $Intention_{t-1}$ • Self-efficacy at the preceding timepoint, expressed as SE_{t-1} • Action planning at the preceding timepoint, expressed as AP_{t-1}	A respondent's HAPA[a] variable score at the timepoint preceding the last available communication score	If an individual had a communication score after module 3, their HAPA variable scores after module 2 were used for the predictor variables.

[a]HAPA: health action process approach.

Multimedia Appendix 2 outlines the intercorrelation between all used social-cognitive HAPA determinants as well as self-reported communication behavior over the course of the app-based intervention. Since the abovementioned analysis includes as many timepoints as possible, all variables at the different timepoints were included.

Results

Study Participants

Overall, 1187 women were recruited into this study, of which 988 were from the Germany-wide recruitment (Figure 1A). Of those in the Germany-wide sample, 506 were randomized into the waitlist control arm (control group app registration), and 482 were randomized into the intervention arm (intervention group app registration). In the clinics, 199 pregnant women were recruited (after the registration of 205 women with the app). The majority of the participants were aged 30-39 years (n=881), had a higher education status (n=763), and were married (n=759).

Descriptive Statistics

Sociodemographic data are depicted in detail in Table 3 below.

XSL•FO
RenderX

Table 3. Sociodemographic characteristics of expectant mothers (N=1187).

Characteristics	Values, n (%)
Age[a] (years)	
18-29	159 (14.23)
30-39	881 (78.87)
40-49	77 (6.89)
Marital status[b]	
Single	31 (2.67)
In a committed relationship	366 (31.5)
Married/registered partnership	759 (65.32)
Divorced/separated/widowed	6 (0.51)
Highest educational level[c]	
No school-leaving qualification	21 (1.81)
Secondary or elementary school leaving	78 (6.74)
Secondary school diploma	137 (11.83)
A-levels	763 (65.89)
Completed vocational training	25 (2.16)
University degree[d]	34 (2.94)
University degree[e]	100 (8.64)

[a]76 missing values for age.
[b]31 missing values for marital status.
[c]35 missing values for highest educational level.
[d]Special German university degree (Hochschule).
[e]University degree.

Predicting Dropout

Of the 1187 pregnant women who were recruited and started using the web-app, 1124 dropped out of the intervention, as indicated by completion of less than 2 modules. A general logistic model was estimated to investigate hypothesis 1 and to determine whether social-cognitive HAPA variables and communication behavior as well as sociodemographic characteristics might predict early dropout (completing less than 2 modules). Thereby, the predictive capacity of 4 HAPA variables and behavior along with age, education, and marital status was tested: intention, outcome expectancy, self-efficacy, and action planning, as well as sex, education, and marital status. As Table 4 highlights, only age was a significant predictor of early dropout. In other words, younger pregnant women were more likely to drop out from the digital intervention at an earlier stage. Accordingly, hypothesis 1 can be empirically supported.

XSL•FO
RenderX

Table 4. Parameter table of the generalized linear model predicting early dropout from health action process approach variables and sociodemographic characteristics.

Variable	Estimate (SE)	*t* test *(df)*[a]	*P* value
Outcome expectancy$_{tt1}$[b]	.005 (.164)	0.034 (418)	.97
Intention$_{tt1}$	–.254 (.153)	–1.658 (418)	.09
Self-efficacy$_{tt1}$	.184 (.121)	1.516 (418)	.13
Action planning$_{t1}$[c]	.204 (.112)	1.816 (418)	.07
Age	–.094 (.038)	–2.469 (418)	.014[d]
Education	–.178 (.161)	–1.106 (418)	.27
Marital status	–.363 (.302)	–1.200 (418)	.23
Recruitment channel=Germany-wide recruitment, randomized into the intervention group (compared to clinical recruitment)	.880 (.523)	1.683 (418)	.09
Recruitment channel=Germany-wide recruitment, randomized into the waitlist control arm (compared to clinical recruitment)	.374 (.540)	0.692 (418)	.49

[a]2-sided *t* test.

[b]tt1: measurement after completing module 1.

[c]t1: module 1 (lessons 1-3).

[d]β is significant at P=.05; R^2=0.13.

Predictors of Self-Reported Communication Behavior

Hypothesis 2 tests whether socio-demographic variables and social-cognitive HAPA variables (self-efficacy, outcome expectancy, and action planning) would predict self-reported communication behavior in pregnant women over the course of the app-based intervention (see Table 5 for details). In the first series of models, each sociodemographic variable was added in a stepwise fashion to predict communication behavior. Adding age ($F_{1,93}$=1.16; P=.28), education ($F_{1,92}$=0.12; P=.66), and family status ($F_{1,91}$=0.39; P=.54) did not significantly improve the prediction of communication behavior scores. In the subsequent models, HAPA variables were added (Table 5), and intention was added. Upon inclusion, most HAPA variables improved the model fit. For the motivational phase of the HAPA model, outcome expectancy ($F_{1,90}$=4.88; P=.03) and intention ($F_{1,98}$=8.65; P=.004) improved the model fit, while task self-efficacy ($F_{1,88}$=2.11; P=.15) did not improve the model fit. For the volitional phase, action planning ($F_{1,87}$=17.74; P<.001) improved the prediction of communication scores.

After including all the variables of HAPA along with sociodemographic variables into the model (for model comparisons, see Table 5), only action planning (β=.331; P<.001) significantly predicted communication behavior (Table 6 for further parameter estimates). Accordingly, hypothesis 2 could be partially empirically supported.

Table 5. Hierarchical regression model comparison of sociodemographic and social-cognitive health action process approach variables predicting communication behavior.

Model name	Comparison model	Predictors	*F* test *(df)*	*P* value
Age	Null	Age	2.165 (1, 92)	.15
Education status	Age	Age + education	0.045 (1, 91)	.83
Marital status	Education status	Age + education + marital status	0.365 (1, 90)	.55
Outcome expectancy	N/A[a]	Outcome expectancy	4.430 (1, 89)	.04[b]
Intention	Outcome expectancy	Outcome expectancy + intention	8.457 (1, 88)	.005[c]
Self-efficacy	Intention	N/A	1.955 (1, 87)	.17
Action planning	Self-efficacy	N/A	17.68 (1, 86)	<.001[d]

[a]N/A: not applicable.

[b]β is significant at P=.05.

[c]β is significant at P=.01.

[d]β is significant at P=.001.

XSL•FO
RenderX

Table 6. Parameter table of hierarchical regression model predicting communication behavior.

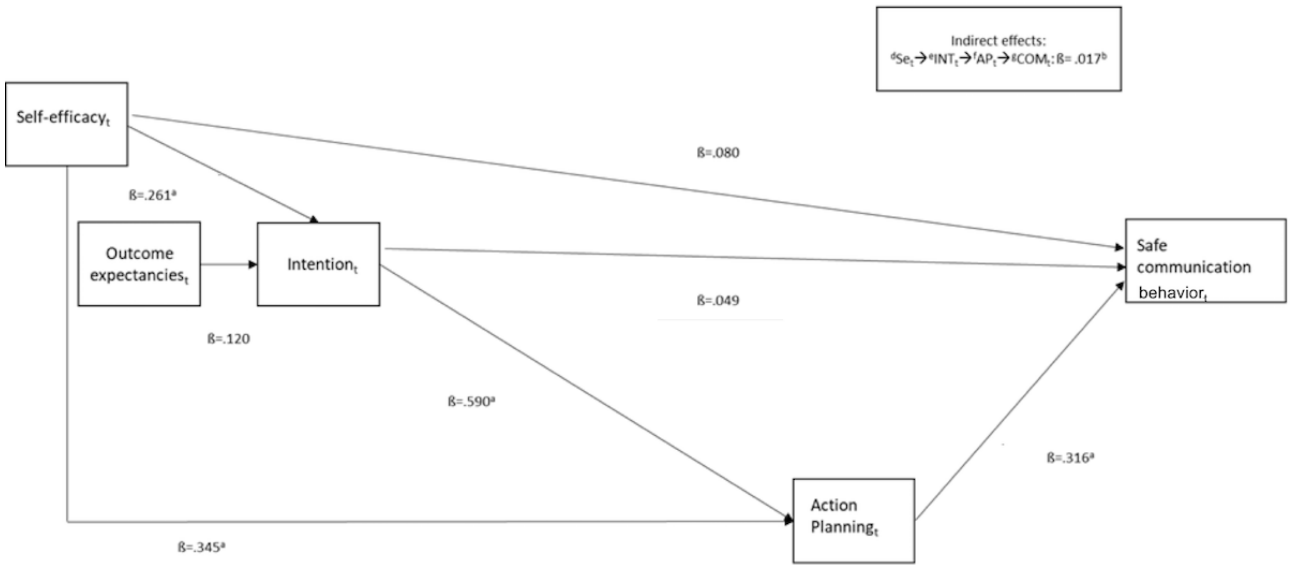
Variable	Estimate (95% CI)	SE	*P* value
Age	–. 016 (–.051 to .018)	.02	.34
Education status	–. 047 (–.154 to .060)	.053	.39
Marital status	–. 100 (–.442 to .243)	.172	.57
Outcome expectancy	.14 (–.047 to .336)	.096	.14
Intention	.046 (–.143 to .236)	.095	.63
Self-efficacy	.007 (–.136 to .149)	.072	.93
Action planning	.305[a] (–.161 to .449)	.072	<.001

[a]β is significant at *P*=.001.

Mediation Model

To test hypothesis 3 and thus whether distal HAPA variables (intention, outcome expectancies, self-efficacy) are mediated through planning, a path model was facilitated (Figure 2 and Table 7). Indeed, a sequential mediation emerged from self-efficacy to intention to action planning to self-reported communication behavior. This likewise entailed the indirect mediation from intention via action planning to self-reported communication behavior. Conversely, no serial mediation was found from outcome expectancies to intention via action planning to self-reported communication behavior. Accordingly, hypothesis 3 could only be empirically supported regarding 2 of the 3 distal HAPA variables, namely, self-efficacy and intention.

Figure 2. Regression model of social-cognitive health action process approach variables and safe communication behavior across all groups. Taken together, communication behavior is significantly predicted by action planning, and action planning mediates the impact of self-efficacy and intention on self-reported communication behavior with β=.017 (comparative fit index=0.994; Tucker–Lewis index=0.971; root mean square error of approximation=0.055). AP: action planning; COM: safe communication behavior; INT: intention; SE: self-efficacy; t: reflects a relative timepoint.



XSL•FO
RenderX

Table 7. Results of the path model depicted in Figure 2.

Predictor variable	Outcome variable	Estimate (95% CI)	SE	*P* value
Self-efficacy	Intention	.353[a] (.152 to .555)	.103	.001
Outcome expectancy	Intention	.105 (–.108 to .318)	.109	.33
Intention	Action planning	.408[a] (.242 to .574)	.085	<.001
Self-efficacy	Planning	.322[a] (.148 to .497)	.089	<.001
Planning	Communication behavior	.471[a] (.311 to .631)	.082	<.001
Intention	Communication behavior	.053 (–.163 to .270)	.111	.63
Self-efficacy	Communication behavior	.108 (–.069 to .285)	.090	.23

[a]β is significant at *P*=.001.

Discussion

Principal Findings

In this study, we investigated the determinants of dropout from an app-based intervention for pregnant women and the mechanisms of adopting self-reported communication behavior. Regarding both aspects, variables from HAPA were measured and evaluated with respect to their predictive capability. Consistent with hypothesis 1, dropout analyses found that only age was predictive, and none of the HAPA variables played a role. Communication behavior was only predicted significantly by one of the HAPA variables, namely, action planning. A serial mediation emerged from intention to self-reported communication behavior via action planning. In detail, communication behavior was significantly predicted by action planning, and action planning mediated the impact of self-efficacy and intention on self-reported communication behavior. These findings match previous findings that identified age as a predictor of dropout from digital interventions [41,42]. However, in contrast to the aforementioned studies, no other sociodemographic predictors emerged. Moreover, this study deviates from previous dropout investigations, as no HAPA variables predicted early dropout. This might be a result of differences in the target group or context of the digital tool between studies; predictors of dropout from digital interventions might depend on aspects of specific intervention types and could vary based on the timepoint of dropout, as revealed in a previous research [63]. Regarding the latter, it should be noted that our study only investigated early dropout.

With respect to mechanisms of adopting self-reported communication behavior, the results were in line with hypotheses 2 and 3: the data demonstrated that HAPA variables predicted the self-reported communication behavior. Whether this is a result of the behavior change context (ie, the app) still needs to be determined. However, the findings are in line with previous evidence that HAPA assumptions match the data and accordingly are able to explain the changes in self-reported communication behavior [8,10].

Our findings of early dropout from the app-based intervention are partially in line with those of other studies, in which younger participants were more likely to drop out from digital interventions [42,64-66]. The relationship between age and dropout could be a result of higher perceived need among older women, that is, older women, through more life experiences and previous pregnancies, may realize a greater need for communication interventions and in turn adhere to the app.

Behavior change variables did not predict dropout from the app-based intervention. This is in contrast to the results of previous studies that have shown that behavior change variables are associated with dropout [63]. Among others, this study shows outcome expectancies as a crucial predictor of retention in digital interventions and likewise concludes that the perception of unmet needs and expectations might be a determinant of dropout [63]. In the context of self-reported communication behavior, expectant mothers' outcome expectancies focused on safe child delivery instead of self-reported communication behavior and the accommodation of individual preferences as the item wording suggests. An alternative explanation would be that the relationship between behavior change variables and dropout is context-specific, as a previous study implicates [63]. In any case, our study shows the determinants of dropout from an app-based intervention. This is an important insight for patient safety interventions in obstetrics because it can be used by future tools to prevent early dropout and maximize the amount of support that pregnant women receive. However, future research should further investigate the contextual variability of predictors of early dropout in digital intervention studies. This is particularly important because app-based interventions generally show high dropout rates [67-70].

Previous studies [8,11] have shown that social-cognitive variables are associated with pregnant women's safe communication behavior in general. Our study demonstrates that some of these associations also drive change in self-reported communication behavior during a digital intervention. This is of importance for theoretical understanding and practitioners; it shows how apps might elicit and affect changes in self-reported communication behavior, highlighting pathways which future interventions can focus on and improve its effectiveness. This might be useful for designing future apps in the specific field of pregnancy and giving birth.

It is striking that not all associations in HAPA emerged as theorized. First, it became apparent that action planning was the single best predictor of change in self-reported communication behavior. Although the predictive capacity of

XSL•FO
RenderX

action planning is expected from its association with behavior within the HAPA model, it was not hypothesized that action planning would emerge as the sole predictor of behavior change. The reason could be pregnant women participating in the app were taught to think of how and when they might communicate effectively, that is, making concrete action plans, which worked well in the app, while other variables were not addressed as effectively. In addition, action planning was targeted in the last lesson of the app, which might have resulted in stronger effects due to a shorter time lag and recency effect.

Relating to the process by which pregnant women improve their communication behavior in clinical care, it is likewise striking why only 2 of the indirect effects specified in the HAPA framework emerged. First, there was an indirect effect of self-efficacy on self-reported communication behavior via intention and subsequently via action planning. Second, self-efficacy also directly impacted action planning and thereby indirectly impacted self-reported communication behavior. Notably, the same predictive capacity in explaining self-reported communication behavior in this study replicated findings from a cross-sectional research [8,11]. Indeed, the HAPA model seems to be applicable for predicting several kinds of behavior change in digitally supported interventions like the app used in this study and has shown similar findings overall [71-73].

In a previous randomized trial in patients with insomnia, both action planning and coping planning in the HAPA model were shown to be effective mediators in improving sleep hygiene [71]. In another randomized trial testing a digital tool to promote active lifestyles in patients with type 2 diabetes, the intervention group showed a significant intervention effect for action planning, whereas the control group exhibited a significant effect for coping planning and self-efficacy [72]. Lastly, in an earlier study primarily focusing on reducing salt intake to prevent high blood pressure, both intention and outcome expectancies as well as risk perception were found to be improved by the digital intervention [73].

In the future, digital and nondigital face-to-face interventions should be compared, especially when aiming to improve self-reported communication behavior in obstetrics and preventing dropout [10]. Different app modes showed various degrees of effectiveness in a study on depression [74]. A meta-analytic review [74] showed that apps in combination with personal contact with a therapist are more effective than self-help apps. However, no differences were found between smartphone-based apps and computer- and internet-based interventions. Similarly, there seems to be no difference between human-guided digital interventions and face-to-face psychotherapy [74]. In other areas of research, gamifications in apps have proven to be beneficial [68]. Among other findings, feedback, leaderboards (participants can compare their own progress with that of others), and storytelling (context within the app to create an alternate reality and guide the user) have been shown to be advantageous for digital interventions [68-70]. Those findings provide some guidance for future teams aiming at developing apps and internet interventions in this field.

Findings from this body of research set the stage for iterating on existing apps in clinical care and for developing new apps. However, it seems questionable, what kind of intervention might be sufficient or helpful for those participants at risk for dropping out. On the one hand, flexible digital tools, which allow an automatic dynamic change of modules and learning intensity, might be helpful. On the other hand, an overload of information or special attention to these participants might make them even more prone to drop out. To conclude, future research should further uncover the reasons for dropout of such participants, so that optimal strategies for prevention can be devised.

Given that younger patients are at increased risk of discontinuing the digital intervention at an early stage, it is crucial to make the underlying behavior visible and targetable. We made the first attempt to explain this phenomenon. In the literature, the possibility of using behavior change theories to predict dropout behavior has been demonstrated [63]. Future research should be conducted using different intervention modes as well as different digital incentives (eg, optimal level of gamification, possibility to exchange with other users vs personal contact with midwives, doctors, or other birthing professionals, or more intensive self-help vs person-guided self-help).

Limitations and Suggestions for Future Research

This study, as a formative research and with preliminary results, has certain limitations with regard to the conclusions drawn from the results. First, there is a possibility that confounding external factors could have been at play during the course of the internet intervention, such as physiological or mental health risk factors. Second, bias and self-selection might have confounded the web-app data in the sense that only certain women volunteered to participate and continue the app-based intervention. Future internet intervention studies should try to recruit a more diverse sample and find concrete reasons for them completing the app or dropping out. It seems possible that both flexible feasibility questions in the context of the app as well as effectiveness ratings and satisfaction ratings could provide more information about usage behavior.

For the time frame between intervention start and the last timepoint, we have conducted a test of factors predicting dropout, from which we concluded that only younger age at intervention start predicted early dropout. However, factors associated with the selection to participate in the intervention could not be uncovered in our presented design. Future research should target a wider age range of pregnant women to gain further insight between age groups and dropout via subgroup analyses. With a larger sample, it also seems possible to examine the age categories and dropout behavior in more detail.

As mentioned previously, our data do not allow conclusions regarding the motivations of dropout and study retainment—a topic that future studies should investigate further. Relatedly, it would have been important to have more finely spaced time intervals for measurement points, which could add valuable information on the interplay of the processes underlying behavioral change and the topics of the particular lessons that were covered. Additionally, it should be acknowledged that the scales used to measure the social-cognitive variables of HAPA were not previous validated in German language based on evaluating the communication behavior. Hence, the

XSL•FO
RenderX

measurement qualities of the scales in the German population might be limited.

The main contribution of this study can be seen in that it is the first attempt of employing a digitally enhanced internet intervention aiming at fostering self-reported communication behavior within a clinical sample in the context of obstetric care. This has innovation potential, as it shows that technology in terms of a digital intervention, that is, apps aiming at improving communication behavior can make a difference. Therefore, this study sheds light on the mechanisms underlying self-reported communication behavior and its improvement while also investigating the potential predictors of dropout in an app-based intervention. This contribution yields both practical and theoretical implications. On a theoretical note, our study contributes to a deeper understanding of the genesis of self-reported communication behavior, thereby highlighting various points of the psychological processes that future interventions could address, such as action planning and self-efficacy. Likewise, practical implications arise as our study presents an initial framework for improving effective communication via the app in a clinical sample and explores how to maximize its effectiveness by retaining participants at risk of early dropout. The variables of the HAPA model can function as a toolkit, with a particular focus on action planning, self-efficacy beliefs, intention to communicate effectively, and app users' age.

Acknowledgments

We thank Katharina Schiffer for proofreading the manuscript regarding referencing style and language. We are thankful and appreciative of the help from Dr Marina Schmiedhofer, Christina Derksen (PhD), Prof Dr Frank Louwen, Prof Dr Frank Reister, Johanna Dietl, Freya Häußler, and Sandra Riedmüller in the data collection. This study was funded by the German Innovation Fund (project 01VSF18023) of the Federal Joint Committee (Gemeinsamer Bundesausschuss) as part of a third-party project funding.

Data Availability

The data that support the findings of this study are available on request from the corresponding author SL. The data are not allowed to be made publicly available due to privacy and data security reasons of the research participants.

Authors' Contributions

LK, VA-K, and SL were involved in data collection and monitoring as well as in the conceptual aspects of this study. LK analyzed and described the data statistically and wrote all parts of this paper. SL advised on the methodology and structure. FMK, VA-K, and NTH gave advice on the rationale and structure of this paper. All coauthors approved the final version and contributed to the preparation as well as revision of the final paper.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Specifications of the TeamBaby web-app.
[DOCX File , 16 KB - humanfactors_v11i1e48218_app1.docx]

Multimedia Appendix 2
Intercorrelation between all used social-cognitive determinants as well as self-reported communication behavior over the course of the app-based intervention.
[DOCX File , 34 KB - humanfactors_v11i1e48218_app2.docx]

References

1. McLaughlin T, Yan Z. Diverse delivery methods and strong psychological benefits: A review of online formative assessment. Computer Assisted Learning 2017 May 25;33(6):562-574. [doi: 10.1111/jcal.12200]
2. Runciman W, Hibbert P, Thomson R, Van Der Schaaf T, Sherman H, Lewalle P. Towards an international classification for patient safety: key concepts and terms. Int J Qual Health Care 2009 Feb;21(1):18-26 [FREE Full text] [doi: 10.1093/intqhc/mzn057] [Medline: 19147597]
3. Eulmesekian PG, Alvarez JP, Ceriani Cernadas JM, Pérez A, Berberis S, Kondratiuk Y. The occurrence of adverse events is associated with increased morbidity and mortality in children admitted to a single pediatric intensive care unit. Eur J Pediatr 2019 Dec 08;179(3):473-482. [doi: 10.1007/s00431-019-03528-z]
4. Schrappe M, Spahn J, Berwick D, Durkin P. APS-Weißbuch. Sicherheit in der Gesundheitsversorgung: Medizinisch Wissenschaftliche Verlagsgesellschaft; 2018.

XSL•FO
RenderX

5. Shojania KG, Marang-van de Mheen PJ. Identifying adverse events: reflections on an imperfect gold standard after 20 years of patient safety research. BMJ Qual Saf 2020 Apr;29(4):265-270. [doi: 10.1136/bmjqs-2019-009731] [Medline: 32066571]
6. Antony J, Zarin W, Pham B, Nincic V, Cardoso R, Ivory JD, et al. Patient safety initiatives in obstetrics: a rapid review. BMJ Open 2018 Jul 06;8(7):e020170 [FREE Full text] [doi: 10.1136/bmjopen-2017-020170] [Medline: 29982200]
7. Gittell JH. Relationships between service providers and their impact on customers. Journal of Service Research 2016 Jun 29;4(4):299-311. [doi: 10.1177/1094670502004004007]
8. Derksen C, Kötting L, Keller FM, Schmiedhofer M, Lippke S. Psychological intervention to improve communication and patient safety in obstetrics: examination of the health action process approach. Front Psychol 2022;13:771626 [FREE Full text] [doi: 10.3389/fpsyg.2022.771626] [Medline: 35250715]
9. Lippke S, Wienert J, Keller FM, Derksen C, Welp A, Kötting L, et al. Communication and patient safety in gynecology and obstetrics - study protocol of an intervention study. BMC Health Serv Res 2019 Nov 28;19(1):908 [FREE Full text] [doi: 10.1186/s12913-019-4579-y] [Medline: 31779620]
10. Derksen C, Dietl JE, Haeussler FE, Steinherr Zazo M, Schmiedhofer M, Lippke S. Behavior change training for pregnant women's communication during birth: A randomized controlled trial. Appl Psychol Health Well Being 2023 Aug;15(3):865-883. [doi: 10.1111/aphw.12413] [Medline: 36380576]
11. Kötting L, Henschel N, Keller FM, Derksen C, Lippke S. Social-cognitive correlates of expectant mothers' safe communication behaviour: Applying an adapted HAPA model. Cogent Psychology 2023 Feb 19;10(1):2173996-2173996. [doi: 10.1080/23311908.2023.2173996]
12. Bertakis KD, Franks P, Azari R. Effects of physician gender on patient satisfaction. J Am Med Womens Assoc (1972) 2003;58(2):69-75. [Medline: 12744418]
13. Dolan JG. Shared decision-making--transferring research into practice: the Analytic Hierarchy Process (AHP). Patient Educ Couns 2008 Dec;73(3):418-425 [FREE Full text] [doi: 10.1016/j.pec.2008.07.032] [Medline: 18760559]
14. Rider EA, Keefer CH. Communication skills competencies: definitions and a teaching toolbox. Med Educ 2006 Jul;40(7):624-629. [doi: 10.1111/j.1365-2929.2006.02500.x] [Medline: 16836534]
15. Attanasio LB, McPherson ME, Kozhimannil KB. Positive childbirth experiences in U.S. hospitals: a mixed methods analysis. Matern Child Health J 2014 Jul;18(5):1280-1290 [FREE Full text] [doi: 10.1007/s10995-013-1363-1] [Medline: 24072597]
16. Delaney AL, Singleton G. Information and relationship functions of communication between pregnant women and their health care providers. Communication Studies 2020 Aug 20;71(5):800-822. [doi: 10.1080/10510974.2020.1807376]
17. Kim SC, Boren D, Solem SL. The Kim Alliance Scale: development and preliminary testing. Clin Nurs Res 2001 Aug;10(3):314-331. [doi: 10.1177/c10n3r7] [Medline: 11881945]
18. Nicoloro-SantaBarbara J, Rosenthal L, Auerbach MV, Kocis C, Busso C, Lobel M. Patient-provider communication, maternal anxiety, and self-care in pregnancy. Soc Sci Med 2017 Oct;190:133-140. [doi: 10.1016/j.socscimed.2017.08.011] [Medline: 28863336]
19. Garnweidner-Holme LM, Lukasse M, Solheim M, Henriksen L. Talking about intimate partner violence in multi-cultural antenatal care: a qualitative study of pregnant women's advice for better communication in South-East Norway. BMC Pregnancy Childbirth 2017 Apr 19;17(1):123 [FREE Full text] [doi: 10.1186/s12884-017-1308-6] [Medline: 28420328]
20. Berry D. Risk, Communication & Health Psychology. UK: McGraw-Hill Education; 2004.
21. Rundell S. A study of nurse-patient interaction in a high dependency unit. Intensive Care Nurs 1991 Sep;7(3):171-178. [doi: 10.1016/0266-612x(91)90006-d] [Medline: 1918904]
22. Yang Y, Zhang X, Lee PK. Improving the effectiveness of online healthcare platforms: An empirical study with multi-period patient-doctor consultation data. International Journal of Production Economics 2019 Jan;207:70-80. [doi: 10.1016/j.ijpe.2018.11.009]
23. Simpson M, Catling C. Understanding psychological traumatic birth experiences: A literature review. Women Birth 2016 Jun;29(3):203-207. [doi: 10.1016/j.wombi.2015.10.009] [Medline: 26563636]
24. Hollander MH, van Hastenberg E, van Dillen J, van Pampus MG, de Miranda E, Stramrood CAI. Preventing traumatic childbirth experiences: 2192 women's perceptions and views. Arch Womens Ment Health 2017 Aug;20(4):515-523. [doi: 10.1007/s00737-017-0729-6] [Medline: 28553692]
25. Lippke S, Derksen C, Keller FM, Kötting L, Schmiedhofer M, Welp A. Effectiveness of communication interventions in obstetrics—a systematic review. IJERPH 2021 Mar 05;18(5):2616. [doi: 10.3390/ijerph18052616]
26. Smith KA, Bishop FL, Dambha-Miller H, Ratnapalan M, Lyness E, Vennik J, et al. Improving empathy in healthcare consultations-a secondary analysis of interventions. J Gen Intern Med 2020 Oct;35(10):3007-3014 [FREE Full text] [doi: 10.1007/s11606-020-05994-w] [Medline: 32666488]
27. D'Agostino TA, Atkinson TM, Latella LE, Rogers M, Morrissey D, DeRosa AP, et al. Promoting patient participation in healthcare interactions through communication skills training: A systematic review. Patient Educ Couns 2017 Jul;100(7):1247-1257 [FREE Full text] [doi: 10.1016/j.pec.2017.02.016] [Medline: 28238421]
28. West R, May S, West M, Croghan E, McEwen A. Performance of English stop smoking services in first 10 years: analysis of service monitoring data. BMJ 2013 Aug 19;347:f4921 [FREE Full text] [doi: 10.1136/bmj.f4921] [Medline: 23963106]

XSL•FO
RenderX

29. Schauer GL, Malarcher AM, Zhang L, Engstrom MC, Zhu S. Prevalence and correlates of quitline awareness and utilization in the United States: an update from the 2009-2010 National Adult Tobacco Survey. Nicotine Tob Res 2014 May;16(5):544-553. [doi: 10.1093/ntr/ntt181] [Medline: 24253378]
30. Moshe I, Terhorst Y, Philippi P, Domhardt M, Cuijpers P, Cristea I, et al. Digital interventions for the treatment of depression: A meta-analytic review. Psychol Bull 2021 Aug;147(8):749-786. [doi: 10.1037/bul0000334] [Medline: 34898233]
31. Dores AR, Geraldo A, Carvalho IP, Barbosa F. The use of new digital information and communication technologies in psychological counseling during the COVID-19 pandemic. Int J Environ Res Public Health 2020 Oct 21;17(20):7663 [FREE Full text] [doi: 10.3390/ijerph17207663] [Medline: 33096650]
32. Rose T, Barker M, Maria Jacob C, Morrison L, Lawrence W, Strömmer S, et al. A systematic review of digital interventions for improving the diet and physical activity behaviors of adolescents. J Adolesc Health 2017 Dec;61(6):669-677 [FREE Full text] [doi: 10.1016/j.jadohealth.2017.05.024] [Medline: 28822682]
33. Kaner EF, Beyer FR, Garnett C, Crane D, Brown J, Muirhead C, et al. Personalised digital interventions for reducing hazardous and harmful alcohol consumption in community-dwelling populations. Cochrane Database Syst Rev 2017 Sep 25;9(9):CD011479 [FREE Full text] [doi: 10.1002/14651858.CD011479.pub2] [Medline: 28944453]
34. Vilardaga R, Casellas-Pujol E, McClernon JF, Garrison KA. Mobile applications for the treatment of tobacco use and dependence. Curr Addict Rep 2019 Jun;6(2):86-97 [FREE Full text] [doi: 10.1007/s40429-019-00248-0] [Medline: 32010548]
35. Marler JD, Fujii CA, Utley DS, Tesfamariam LJ, Galanko JA, Patrick H. Initial assessment of a comprehensive digital smoking cessation program that incorporates a mobile app, breath sensor, and coaching: cohort study. JMIR Mhealth Uhealth 2019 Feb 04;7(2):e12609 [FREE Full text] [doi: 10.2196/12609] [Medline: 30670372]
36. Whittaker R, McRobbie H, Bullen C, Rodgers A, Gu Y, Dobson R. Mobile phone text messaging and app-based interventions for smoking cessation. Cochrane Database Syst Rev. (CD006611). PMC680 2019;10:4292. [doi: 10.1002/14651858.cd006611.pub5]
37. Marcu G, Ondersma SJ, Spiller AN, Broderick BM, Kadri R, Buis LR. The perceived benefits of digital interventions for behavioral health: qualitative interview study. J Med Internet Res 2022 Mar 30;24(3):e34300 [FREE Full text] [doi: 10.2196/34300] [Medline: 35353045]
38. Kripalani S, LeFevre F, Phillips CO, Williams MV, Basaviah P, Baker DW. Deficits in communication and information transfer between hospital-based and primary care physicians: implications for patient safety and continuity of care. JAMA 2007 Feb 28;297(8):831-841. [doi: 10.1001/jama.297.8.831] [Medline: 17327525]
39. Karin E, Dear BF, Heller GZ, Crane MF, Titov N. "Wish You Were Here": examining characteristics, outcomes, and statistical solutions for missing cases in web-based psychotherapeutic trials. JMIR Ment Health 2018 Apr 19;5(2):e22 [FREE Full text] [doi: 10.2196/mental.8363] [Medline: 29674311]
40. Torous J, Lipschitz J, Ng M, Firth J. Dropout rates in clinical trials of smartphone apps for depressive symptoms: A systematic review and meta-analysis. J Affect Disord 2020 Feb 15;263:413-419. [doi: 10.1016/j.jad.2019.11.167] [Medline: 31969272]
41. Karyotaki E, Kleiboer A, Smit F, Turner DT, Pastor AM, Andersson G, et al. Predictors of treatment dropout in self-guided web-based interventions for depression: an 'individual patient data' meta-analysis. Psychol Med 2015 Oct;45(13):2717-2726 [FREE Full text] [doi: 10.1017/S0033291715000665] [Medline: 25881626]
42. Moshe I, Terhorst Y, Paganini S, Schlicker S, Pulkki-Råback L, Baumeister H, et al. Predictors of dropout in a digital intervention for the prevention and treatment of depression in patients with chronic back pain: secondary analysis of two randomized controlled trials. J Med Internet Res 2022 Aug 30;24(8):e38261 [FREE Full text] [doi: 10.2196/38261] [Medline: 36040780]
43. Wu MS, Chen S, Wickham RE, Leykin Y, Varra A, Chen C, et al. Predicting non-initiation of care and dropout in a blended care CBT intervention: Impact of early digital engagement, sociodemographic, and clinical factors. Digit Health 2022;8:20552076221133760 [FREE Full text] [doi: 10.1177/20552076221133760] [Medline: 36312847]
44. Gao L, Keller FM, Becker P, Dahmen A, Lippke S. Predictors of dropout among psychosomatic rehabilitation patients during the COVID-19 pandemic: secondary analysis of a longitudinal study of digital training. J Med Internet Res 2023 Nov 27;25:e43584 [FREE Full text] [doi: 10.2196/43584] [Medline: 37903289]
45. Cerdà-Navarro A, Quintana-Murci E, Salvà-Mut F. Reasons for dropping out of intermediate vocational education and training in Spain: the influence of sociodemographic characteristics and academic background. Journal of Vocational Education & Training 2022 Mar 06:1-25. [doi: 10.1080/13636820.2022.2049625]
46. Davis MJ, Addis ME. Predictors of attrition from behavioral medicine treatments. Ann Behav Med 1999;21(4):339-349. [doi: 10.1007/BF02895967] [Medline: 10721442]
47. Vöhringer M, Knaevelsrud C, Wagner B, Slotta M, Schmidt A, Stammel N, et al. Should I stay or must I go? Predictors of dropout in an internet-based psychotherapy programme for posttraumatic stress disorder in Arabic. Eur J Psychotraumatol 2020;11(1):1706297 [FREE Full text] [doi: 10.1080/20008198.2019.1706297] [Medline: 32082510]
48. Murray E, White IR, Varagunam M, Godfrey C, Khadjesari Z, McCambridge J. Attrition revisited: adherence and retention in a web-based alcohol trial. J Med Internet Res 2013 Aug 30;15(8):e162 [FREE Full text] [doi: 10.2196/jmir.2336] [Medline: 23996958]

XSL•FO
RenderX

49. Schroé H, Crombez G, De Bourdeaudhuij I, Van Dyck D. Investigating when, which, and why users stop using a digital health intervention to promote an active lifestyle: secondary analysis with a focus on health action process approach-based psychological determinants. JMIR Mhealth Uhealth 2022 Jan 31;10(1):e30583 [FREE Full text] [doi: 10.2196/30583] [Medline: 35099400]
50. Maenhout L, Peuters C, Cardon G, Crombez G, DeSmet A, Compernolle S. Nonusage attrition of adolescents in an mHealth promotion intervention and the role of socioeconomic status: secondary analysis of a 2-arm cluster-controlled trial. JMIR Mhealth Uhealth 2022 May 10;10(5):e36404 [FREE Full text] [doi: 10.2196/36404] [Medline: 35536640]
51. Donkin L, Glozier N. Motivators and motivations to persist with online psychological interventions: a qualitative study of treatment completers. J Med Internet Res 2012 Jun 22;14(3):e91 [FREE Full text] [doi: 10.2196/jmir.2100] [Medline: 22743581]
52. Davis R, Campbell R, Hildon Z, Hobbs L, Michie S. Theories of behaviour and behaviour change across the social and behavioural sciences: a scoping review. Health Psychol Rev 2015;9(3):323-344 [FREE Full text] [doi: 10.1080/17437199.2014.941722] [Medline: 25104107]
53. Dillon K, Rollo S, Prapavessis H. A combined health action process approach and mHealth intervention to reduce sedentary behaviour in university students - a randomized controlled trial. Psychol Health 2022 Jun;37(6):692-711. [doi: 10.1080/08870446.2021.1900574] [Medline: 33780297]
54. Schwarzer R. Modeling health behavior change: how to predict and modify the adoption and maintenance of health behaviors. Applied Psychology 2008 Jan 30;57(1):1-29. [doi: 10.1111/j.1464-0597.2007.00325.x]
55. Every newborn: an action plan to end preventable deaths. WHO. 2014. URL: https://www.who.int/initiatives/every-newborn-action-plan [accessed 2023-01-03]
56. Schwarzer R, Luszczynska A. Self-efficacy. Division of Cancer Control and Population Sciences (DCCPS). URL: https://cancercontrol.cancer.gov/brp/research/constructs/self-efficacy [accessed 2023-02-16]
57. Schwarzer R, Hamilton K. Changing behavior using the Health Action Process Approach. In: The Handbook of Behavior Change. Cambridge, England: Cambridge University Press; 2020:89-103.
58. Keller FM, Derksen C, Kötting L, Schmiedhofer M, Lippke S. Development of the perceptions of preventable adverse events assessment tool (PPAEAT): measurement properties and patients' mental health status. Int J Qual Health Care 2021 Apr 16;33(2):mzab063. [doi: 10.1093/intqhc/mzab063] [Medline: 33822086]
59. Gholami M, Schwarzer R, Reyes-Fernandez B, Rica C. Brief scales for the multilingual assessment of HAPA variables. Health Psychology. URL: https://www.psyc.de/hapascales.pdf [accessed 2023-01-10]
60. Accessed April 3, 2023. R: The R Project for Statistical Computing. URL: https://www.r-project.org/ [accessed 2023-01-05]
61. RStudio T. RStudio: integrated development for R. RStudio. URL: http://www.rstudio.com/ [accessed 2023-01-02]
62. Rosseel Y. lavaan: An R package for structural equation modeling. Journal of Statistical Software 2012;48:1-36. [doi: 10.18637/jss.v048.i02]
63. Lippke S, Gao L, Keller FM, Becker P, Dahmen A. Adherence with online therapy vs face-to-face therapy and with online therapy vs care as usual: secondary analysis of two randomized controlled trials. J Med Internet Res 2021 Nov 03;23(11):e31274 [FREE Full text] [doi: 10.2196/31274] [Medline: 34730541]
64. Linnet J, Hertz SPT, Jensen ES, Runge E, Tarp KHH, Holmberg TT, et al. Days between sessions predict attrition in text-based internet intervention of binge eating disorder. Internet Interv 2023 Mar;31:100607 [FREE Full text] [doi: 10.1016/j.invent.2023.100607] [Medline: 36819741]
65. Jiandani D, Wharton S, Rotondi MA, Ardern CI, Kuk JL. Predictors of early attrition and successful weight loss in patients attending an obesity management program. BMC Obes 2016;3:14 [FREE Full text] [doi: 10.1186/s40608-016-0098-0] [Medline: 26966544]
66. Van der Mispel C, Poppe L, Crombez G, Verloigne M, De Bourdeaudhuij I. A self-regulation-based eHealth intervention to promote a healthy lifestyle: investigating user and website characteristics related to attrition. J Med Internet Res 2017 Jul 11;19(7):e241 [FREE Full text] [doi: 10.2196/jmir.7277] [Medline: 28698168]
67. Eysenbach G. The law of attrition. J Med Internet Res 2005 Mar 31;7(1):e11 [FREE Full text] [doi: 10.2196/jmir.7.1.e11] [Medline: 15829473]
68. Christensen H, Griffiths KM, Farrer L. Adherence in internet interventions for anxiety and depression. J Med Internet Res 2009 Apr 24;11(2):e13 [FREE Full text] [doi: 10.2196/jmir.1194] [Medline: 19403466]
69. Hernández-Rodríguez JC, García-Muñoz C, Ortiz-Álvarez J, Saigí-Rubió F, Conejo-Mir J, Pereyra-Rodriguez J. Dropout rate in digital health interventions for the prevention of skin cancer: systematic review, meta-analysis, and metaregression. J Med Internet Res 2022 Dec 09;24(12):e42397 [FREE Full text] [doi: 10.2196/42397] [Medline: 36485027]
70. Peels DA, Bolman C, Golsteijn RHJ, De Vries H, Mudde AN, van Stralen MM, et al. Differences in reach and attrition between Web-based and print-delivered tailored interventions among adults over 50 years of age: clustered randomized trial. J Med Internet Res 2012 Dec 17;14(6):e179 [FREE Full text] [doi: 10.2196/jmir.2229] [Medline: 23246790]
71. Rajabi Majd N, Broström A, Ulander M, Lin C, Griffiths MD, Imani V, et al. Efficacy of a theory-based cognitive behavioral technique app-based intervention for patients with insomnia: randomized controlled trial. J Med Internet Res 2020 Apr 01;22(4):e15841 [FREE Full text] [doi: 10.2196/15841] [Medline: 32234700]

XSL•FO
RenderX

72. Poppe L, De Bourdeaudhuij I, Verloigne M, Shadid S, Van Cauwenberg J, Compernolle S, et al. Efficacy of a self-regulation-based electronic and mobile health intervention targeting an active lifestyle in adults having type 2 diabetes and in adults aged 50 years or older: two randomized controlled trials. J Med Internet Res 2019 Aug 02;21(8):e13363 [FREE Full text] [doi: 10.2196/13363] [Medline: 31376274]
73. Fuchs K, Meusburger D, Haldimann M, Ilic A. NutritionAvatar: designing a future-self avatar for promotion of balanced, low-sodium diet intention: framework design and user study. 2019 Presented at: Proceedings of the 13th Biannual Conference of the Italian SIGCHI; September; Italy p. 1-10. [doi: 10.1145/3351995.3352044]
74. Moshe I, Terhorst Y, Philippi P, Domhardt M, Cuijpers P, Cristea I, et al. Digital interventions for the treatment of depression: A meta-analytic review. Psychol Bull 2021 Aug;147(8):749-786. [doi: 10.1037/bul0000334] [Medline: 34898233]

Abbreviations

HAPA: health action process approach

Edited by A Kushniruk; submitted 16.04.23; peer-reviewed by A Hannawa, W Zhang, M Gasmi ; comments to author 10.10.23; revised version received 31.01.24; accepted 20.02.24; published 26.04.24.

Please cite as:
Kötting L, Anand-Kumar V, Keller FM, Henschel NT, Lippke S
Effective Communication Supported by an App for Pregnant Women: Quantitative Longitudinal Study
JMIR Hum Factors 2024;11:e48218
URL: https://humanfactors.jmir.org/2024/1/e48218
doi:10.2196/48218
PMID:38669073

©Lukas Kötting, Vinayak Anand-Kumar, Franziska Maria Keller, Nils Tobias Henschel, Sonia Lippke. Originally published in JMIR Human Factors (https://humanfactors.jmir.org), 26.04.2024. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on https://humanfactors.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Comparing Attitudes Toward Different Consent Mediums: Semistructured Qualitative Study

Xengie Doan[1], MS; Arianna Rossi[2], PhD; Marietjie Botes[1,3], LLD; Annika Selzer[4,5], PhD

[1]SnT, University of Luxembourg, Esch-sur-Alzette, Luxembourg
[2]LIDER Lab, DIRPOLIS, Scuola Superiore Sant'Anna, Pisa, Italy
[3]Department of Medicine, Stellenbosch University, Stellenbosch, South Africa
[4]Fraunhofer Institute for Secure Information Technology, Darmstadt, Germany
[5]ATHENE, National Research Center for Applied Cybersecurity, Darmstadt, Germany

Corresponding Author:
Xengie Doan, MS
SnT
University of Luxembourg
2 place de l'Université
Esch-sur-Alzette, 4365
Luxembourg
Phone: 352 46 66 44 1
Email: xengie.doan@uni.lu

Abstract

Background: As consent for data sharing evolves with the digital age, plain-text consent is not the only format in which information can be presented. However, designing a good consent form is highly challenging. The addition of graphics, video, and other mediums to use can vary widely in effectiveness; and improper use can be detrimental to users.

Objective: This study aims to explore the expectations and experiences of adults toward consent given in infographic, video, text, newsletter, and comic forms in a health data sharing scenario to better understand the appropriateness of different mediums and identify elements of each medium that most affect engagement with the content.

Methods: We designed mock consent forms in infographic, video, text, newsletter, and comic versions. Semistructured interviews were conducted with adults who were interviewed about their expectations for consent and were then shown each consent medium and asked about engaging elements across mediums, preferences for consent mediums, and the value of document quality criteria. We transcribed and qualitatively co-coded to identify themes and perform analyses.

Results: We interviewed 24 users and identified different thematic archetypes based on participant goals, such as the Trust Seeker, who considered their own understanding and trust in organizations when making decisions. The infographic was ranked first for enhancing understanding, prioritizing information, and maintaining the proper audience fit for serious consent in health data sharing scenarios. In addition, specific elements such as structure, step-by-step organization, and readability were preferred engaging elements.

Conclusions: We identified archetypes to better understand user needs and elements that can be targeted to enhance user engagement with consent forms; this can help inform the design of more effective consent in the future. Overall, preferences for mediums are highly contextual, and more research should be done.

(JMIR Hum Factors 2024;11:e53113) doi:10.2196/53113

KEYWORDS

consent; transparency; data governance; visualization; health data sharing

Introduction

Overview

Consent is a cornerstone of ethical research, allowing people to be informed about the risks and benefits of research and demonstrate their autonomy. Consent has been discussed since the Nuremberg trials and takes on a pivotal role throughout European Union (EU) regulations for data protection, such as the General Data Protection Regulation (GDPR), but there are still challenges as bioethical consent and data protection consent

XSL•FO
RenderX

collide. Digital decision-making about one's own data can be influenced or misled through interface design choices (ie, through so-called dark patterns [1-8]), while the consent experience of most European users corresponds to nagging cookie consent requests with profiling and advertisements that induce consent fatigue while trying to access a needed service [8]. Decades of research in the biomedical domain show that study participants' consent can rarely be deemed actually informed [9], often due to the complexity of language [10] and lack of health literacy [11], as well as the lack of data literacy of the individuals [12].

Engaging individuals in a user-friendly consent experience is thus fundamental to enabling them to meaningfully and freely make decisions with a sense of satisfaction [10] and agency. Improving the readability and comprehensibility of consent notices is one aspect of this, but research is also being done to explore visual communication techniques. Current research often focuses on the effect of multimedia on understanding [11,12], which can have a varied effect based on different studies. Multimedia also spans many formats, and most studies reviewed for their effect on understanding compared 2-3 different formats [12]. The Article 29 Working Party also refers to visual design means, such as "cartoons, infographics, flowcharts," to enhance the comprehensibility of information, and specifically to "comics/cartoons, pictograms, animations" [13]. However, they do not offer further guidance about what mediums to use and for what purpose (eg, how one might prioritize skimming, while another might be better for complex information). Therefore, we experimented with 5 different mediums of consent in this study, building on studies researching the use of a comic [14,15], video [16], infographic, and illustrated text [17], with plain text as a control [18].

In this paper, we substantially built on our previous work [19] by analyzing more mediums beyond the comic and infographic and specific engaging elements. The study presented a fictional scenario with a data trustee who would assist organizations (eg, research institutions, hospitals, etc) in finding suitable participants for clinical trials in a privacy-friendly manner. Participants were given a scenario where they were individuals who may benefit from a clinical trial organized by a hospital, so the data intermediary requested their consent to share their contact information with the hospital.

The objective of this study was to better profile user expectations and their attitudes toward different consent mediums, which included infographic, video, text, newsletter, and comic versions. We specifically analyzed how different elements of consent mediums (eg, narrative, color, and audio) affected participant engagement to survey the different affordances of each medium. Each medium has its own strengths and weaknesses in representing various kinds of information and can achieve various informational goals (eg, the video is low effort but can be skimmed, while the text can be skimmed but boring) [20]. As we intended to understand whether there are benefits to using one medium over another and why participants would prefer different mediums, we compared multiple mediums in this study based on semistructured interviews and dived into participant motivations, expectations, and experiences.

The results hinted at diverse goals among participants. We also identified the elements of document design that make the information concise, structured, and appropriate for the audience. We also found a large influence of context (eg, cookie consent or consent with different trusted institutions) on participant perceptions and expectations. Thus, we offer recommendations on how to better design consent documents to address different general participant profiles using layering and to engage the audience more effectively with a suitable medium. This has a pivotal role in the digital health data sharing space to give more effective transparency to participants who are deciding whether to share sensitive data. Our results can be leveraged by designers of digital consent experiences for more efficient multimedia use.

Background

Consent and Transparency

The European data strategy [21,22] aims to create a single market for data to allow for the free flow of data to benefit businesses, research, and public administrations within the EU. It is built on the GDPR, which aims to give users more control over their personal data.

Informed consent (IC) is a legal requirement specified in the GDPR as "freely given, specific, informed and unambiguous" (article 4(11)); easily withdrawn (article 7(3)); presented in an intelligible and easily accessible form using clear and plain language (article 7(2)); explicitly given for biomedical and genome data categorized as sensitive data (article 9); transparent in terms of completeness, comprehensibility, and accessibility of the information disclosures (articles 12, 13, and 14); and compliant with the principles of data protection by design and by default (article 25) [23]. IC requires user-centric design elements in consent to help achieve the general principle of transparency, which encompasses the "quality, accessibility, and comprehensibility of the information" [14]. The GDPR also contains obligations for "transparency by design" wherein privacy and consent notices should be purposefully designed to adequately inform the intended audience [24]. In addition, the GDPR also refers to other visual design methods like comics, videos, and infographics.

However, most existing informed decision-making solutions fail to reconcile theoretical demands with actual transparency. Conventional data privacy communication is characterized by lengthy, off-putting walls of complex jargon that impact the readability, comprehensibility, navigability, and memorability of information [20]. In addition, it is often standard, vague, or boilerplate instead of customized to the different needs and abilities of the intended audiences [25] and the type of data and processing activity. Reaching beyond plain language, in the last few years, there has been a renewed attention (and quite some experimentation) toward legal document design criteria [26] that more holistically relate to the language, writer-reader relationship, information design, and content.

Profiling User Needs Using Archetypes

Human-computer interaction research has used the persona technique (wherein imaginary users are assigned different profiles or personas with different goals and personalities based

XSL•FO
RenderX

on demographic data) to better understand different users and needs and design suitable solutions [27]. However, it is a lengthy process that is often used for designing IT systems, not the consent processes. User need assessments have been conducted in relation to different demographics in health studies, but rather than focusing on the IC aspect, they focus on the health symptoms and how to address specific health-related needs [18,28,29]. Beyond health-related needs, we are interested more broadly in how the general adult population would interact with consent process to share information for downstream health reasons and what elements would be engaging when making informed decisions. This aspect has not been studied, to the best of our knowledge, but would be important for understanding how to strategically create effective information disclosures. Thus, we wanted to create archetypes, which capture general profiles, instead of personas, which are representations of imaginary individuals with specific population characteristics.

Multimedia Tools and Engagement With Digital Consent

The digitalization of data collection and use authorization allows for multimedia tools to be used during the consent process, which can have a positive outcome for participants. Overall, a systematic review of multimedia consent with videos, interactive programs, so on for surgical procedures found increased patient satisfaction for usability and informational availability [30]. However, for clinical trial consent, videos did not improve understanding [31]. Diving into the reasons that multimedia consent may be preferred to conventional text, one study compared animated videos, slideshows with voice-overs, comics, and text consent for medical practices and found that a dual-channel approach combining audio with visuals helped participant understanding [31]. This study supported older research showing that repetition of information using different multimedia means increases retention [32]. However, the specific elements of videos, comics, and text that contributed to effective communication in more general health consent were not studied—a gap that we intend to bridge with our work.

Even in other domains, studies strive to understand how to achieve effective communication of complex information by analyzing participant engagement, understanding, and recall of the information [33]. In the study by Wang et al [33], engagement refers to the time spent and fun experienced reading a form; and infographics, illustrated text, and data comics of complex economic data were tested. They found that students from different countries (aged from 18 to 35 years) preferred data comics, as they enable the greatest understanding, engagement, and enjoyment of all mediums, while the infographic performed best in esthetics and exploration, and the illustrated text performed the worst. As similar studies had not been performed on consent forms in a health scenario, we sought to study engagement as a factor of effective communication, as it might help understand what gains and retains attention within a complex digital attention economy.

Traditionally, engagement studies in biomedical consent refer to patient engagement with the research or biomedical process. Such engagement refers to participants interacting with the results of a study, updating information, or changing consent [34-36]. However, we are interested in participant motivations to consume the information in a consent form and give their initial and continued attention to a conventionally tedious process while competing in an attention economy [37,38]. Can consent forms be interesting and attention-grabbing?

Research Questions

The previous section has gathered evidence about the interplay between GDPR transparency requirements in data protection, the use of archetypes, and multimedia tools to enhance the experience. However, we lack an understanding of user needs, the impacts of different mediums on the user experience, and user engagement. The research questions that this study sought to answer are shown in Textbox 1.

Textbox 1. Research questions.

1. What kind of goal-oriented archetypes can be created to better understand participant needs for consent?
2. Across the 5 analyzed mediums (ie, infographic, video, text, newsletter, and comic forms), what were the participant rankings of different engaging elements?
3. After exposure to the consent mediums, we asked the following questions:
    a. What were the participants' rankings of consent mediums?
    b. What elements reportedly influence their preference for mediums?
    c. What document quality criteria concerning language, design, content, and relationship with the reader did participants value?

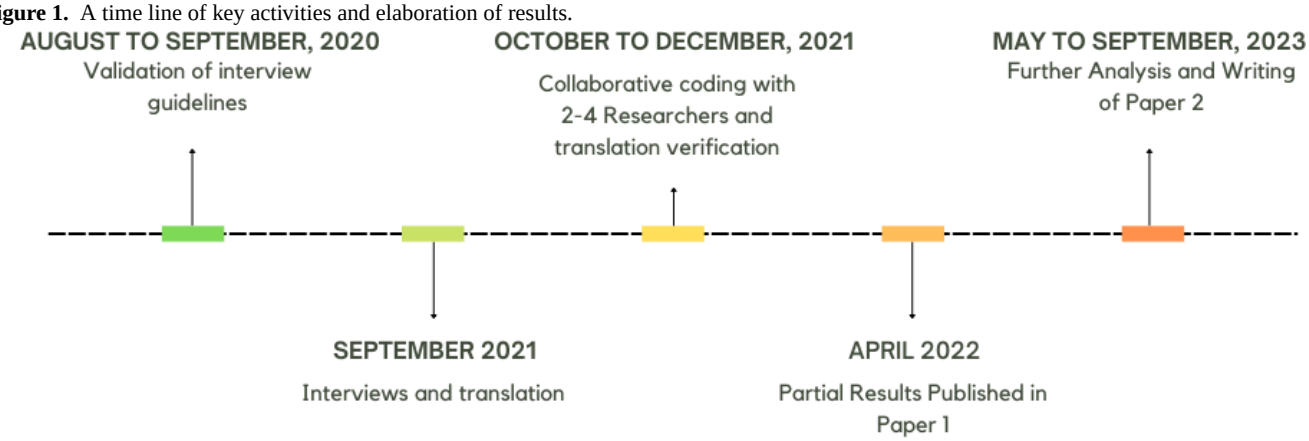
Methods

Overview

AS carried out 24 semistructured interviews in September 2021 in Germany (Figure 1). We created an interview guideline (Multimedia Appendix 1), which was validated with 3 potential participants to ensure clarity, comprehensibility, and precision of the questions.

XSL•FO
RenderX

Figure 1. A time line of key activities and elaboration of results.



Recruitment

We searched for 24 participants by word of mouth. The demographic included adults from a cross-section of the German adult population by age, sex, and education level (Table 1).

Table 1. Participant demographics (N=24).

Characteristics	Participants, n (%)
Age range (y)	
18-30	8 (33)
31-55	8 (33)
56-90	8 (33)
Sex	
Male	12 (50)
Female	12 (50)
Highest degree	
School leaving or apprenticeship	12 (50)
College or university	12 (50)

The sample size and participant characteristics were based on a systemic review of unbiased citizens' juries for health policies [39]. Within the age ranges, there was an equal distribution of men and women with the highest degree obtained. All participants were native German speakers and lived in Germany, and the interviews took an average of 60 to 75 minutes.

Study Material

AS created an example plain text document that asked for consent for the transfer of personal data from an intermediation service to another organization, a hospital. On the basis of the plain text, XD designed 4 additional variations in different mediums: an infographic, a comic, a newsletter, and a video. These 4 variations only included the subsection "What happens if you agree?" of the consent form. All 5 consent forms (ie, plain text, infographic, comic, newsletter, and video forms) differed in design, but the core consent text was the same across all mediums. XD followed best practices for information transparency, designed documents for each medium with different subsets of engaging elements, and adapted them for the mediums (ie, additional ellipses between comic text) for the purposes of the study, consulting coauthors during design. (a) The video (Multimedia Appendix 2) was created to test the use of color, audio, and animations and illustrate the text using free resources on Biteable website (eg, "a doctor will call you" conveyed as an animation of a waving physician). The audio was provided by AS (a native German speaker) out of convenience. (b) The infographic (Figure 2) was designed with a step-by-step format and color from a health template on Canva (Canva Pty Ltd), with icons describing the text (eg, scheduling an appointment had a calendar icon; Multimedia Appendix 3). (c) The comic (Figure 3) used a story element and color and was designed in Figma with input from all authors, and it used simple figures to expedite the creation of the comics. The drawings sought to describe the text as literally as possible (eg, "you will be contacted" depicts a ringing phone; Multimedia Appendix 4). (d) The newsletter (Figure 4) used open format and color and was created in Figma (Figma, Inc; a popular website for user interface or user experience design) based on an existing newsletter template's structure. The newsletter was thought to be a more familiar medium with more graphics than text (eg, newsletters sent via email; Multimedia Appendix 5).

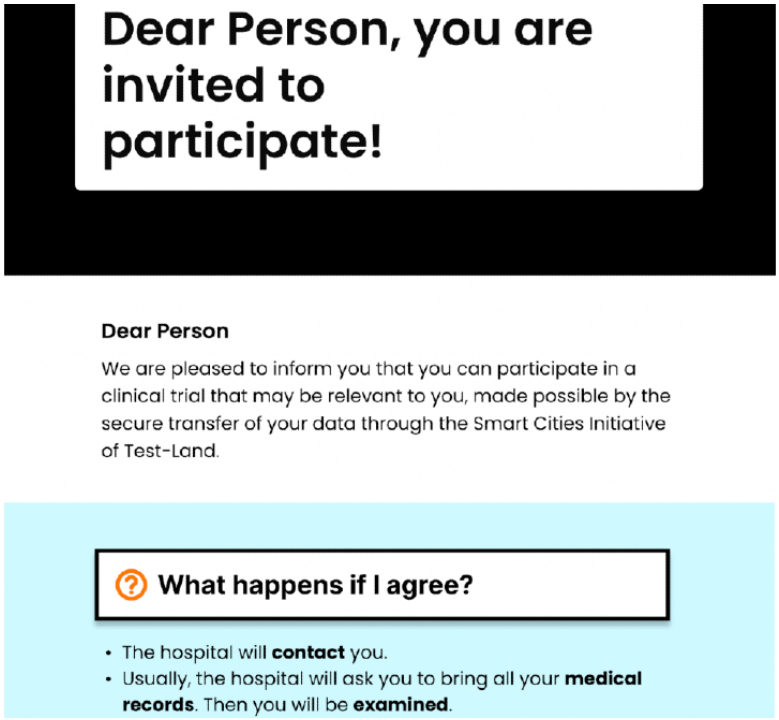
XSL•FO
RenderX

Figure 2. A translated section of the infographic study material designed with a step-by-step format, color, and structured sections.

Figure 3. A translated section of the comic study material designed with a story and color.

XSL•FO
RenderX

Figure 4. A translated section of the newsletter study material designed with an open format, color, and structured sections.



Study Design

All interviews took place via a web conferencing system. No recordings were taken, and a summary transcription was written after each question and finalized after each interview. This method was chosen to respect participant anonymity and COVID-19 protocols.

The interviewer invited participants to imagine that they were contacted by a data intermediary to obtain their consent to share their name, email, and allergy information with a hospital that wanted to carry out a clinical trial for lactose allergies. A verbal explanation was given along with the full plain text version of the consent form, and participants could ask questions at any time.

To answer RQ1, participants were asked about their previous experience with consent forms and desires regarding consent.

To address RQ2, after participants were shown all mediums of consent, they were asked to rank 8 design elements of a consent form: the use of colors, audio, animated elements, readability of text (eg, if it is not too technical or complicated), story element (eg, using examples and people in the forms), structured sections, step-by-step elements (eg, having an order to the information with text or visuals), and an open format (eg, being able to skip around to sections) from the most to the least engaging with an option for "other."

To answer RQ3, we showed them a subsection of the full consent form, "What happens if I agree?" in different mediums (ie, comic, infographic, plain text, newsletter, and video versions) in a random order per participant. Participants were asked to rank the different forms according to their preferences and clarify why.

Data Analysis

The interviews were documented in German, and anonymized answers were translated into English via DeepL (DeepL SE) and proofread by AS to ensure the translations' adherence to the original meaning and to collaboratively analyze them with XD, AR, and MB (all non-German). Translation verification continued throughout the qualitative coding process in various sessions from November 2021 to April 2022 with the multidisciplinary team. To code the interview, the software MAXQDA (VERBI GmbH) was used. The expertise of the coding team spanned data protection law, usable privacy, bioethics, bioinformatics, and legal design.

To code the interviews, we inductively and iteratively established a codebook over three 2-hour sessions (Multimedia Appendix 6). The codebook combines a bottom-up approach through analysis of the data (eg, the concept of trust stemming from participant answers) with a top-down approach derived from the criteria for good documents given by Waller [26] (Table 2) to answer RQ3 (c).

XSL•FO
RenderX

Table 2. Document quality criteria elaborated by Waller [26].

Criteria	Description
Language	
Directness	Using direct language to make it clear who is acting
Plain words	The extent to which the vocabulary is easily understood
Grammar	Conformity with good standard English practices
Readability	Ease with which the reader can follow arguments
Design	
Legibility	Use of legible fonts and text layout
Graphic elements	Use of tables, bullet lists, graphs, charts, icons, etc
Structure	Quality of document organization for function
Impression	Attractiveness and approachability, overall appearance
Relationship	
Who from	Is it clear who is communicating?
Contact	Whether there are clear contacts or means of contact
Audience fit	Appropriateness to the knowledge and skills of users
Tone	Matching the style and language of the context
Content	
Relevance	How relevant is the content to the recipient?
Subject	If it is clear what the communication is about
Action	Clarity about what action is required of the user
Alignment	Compliance with the organization's intended aims and values

Participant consent expectations have been organized into archetypes depending on the salience of reported goals and relevant features. A matrix was created with the participant number, expected features, expected goals, and expected behaviors to help group similar profiles.

Ethical Considerations

The study design has been authorized by the Research Ethics Committee of the University of Luxembourg (ERP 21-038 LeAds), and best practices were followed. We chose a summary transcription to enable easier anonymization of the interview. Once manually anonymized, transcripts were securely shared with the authors from the other organization. The interviewees were compensated €30 for their time.

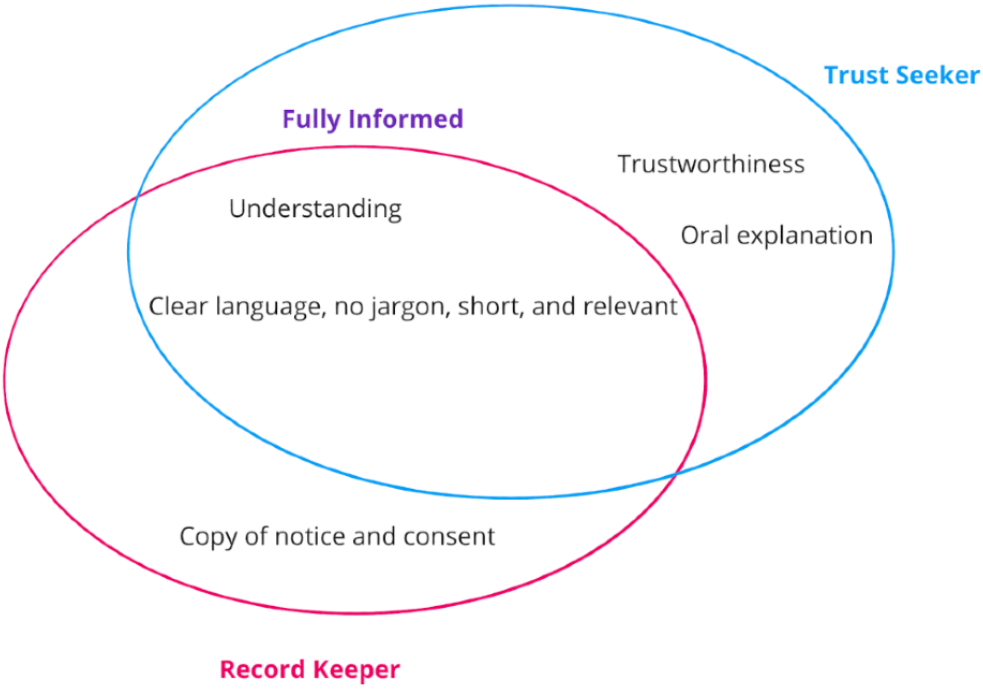
Results

User Desires via Archetypes

The interview findings from the questions, which explore participant expectations, desires, and needs, have been organized into 3 goal-oriented archetypes: the Fully Informed, the Record Keeper, and the Trust Seeker. Not all participants reported specific goals, while some participants reported multiple goals. Thus, the archetypes are based on grouping similar features (Figure 5).

XSL•FO
RenderX

Figure 5. Venn diagram describing the core goals of different consent archetypes.



The Fully Informed archetype wanted relevant and fitting information to understand what they were consenting to. This aligns with the most common goal explicitly reported by participants (14/24, 58%):

> *[A]s an affected person, I would like to see a few examples to get a better understanding of what may be done with my data.* [P1]

The information must also be appropriate for them as an audience:

> *A simple explanation that everyone understands would be my preference.* [P12]

The Record Keeper sought understanding while specifically wanting to remember what they had agreed to (3/24, 13%) or to have a copy for their records (4/24, 17%). For example, participant 13 had a clear idea of the elements they wanted to understand and retain a clear memory of:

> *It needs to be clear to me what the consent is for, who it is from, and exactly what data is being processed for what purpose.* [P13]

In addition, participant 4 stated the following:

> *It doesn't matter to me if it is paper or digital. The main thing is that I receive a copy of the text to which I have consented.* [P4]

The Trust Seeker also sought understanding but was cautious toward the system or desired a trustworthy system (6/24, 25%):

> *I must have the impression that the data trustee is a reliable company or that there is an expertise that proves that I can trust this data trustee.* [P3]
>
> *[I would rather avoid] to invest time and read through stuff...[and be able to trust] since I've already given my data...that my data will just be handled well.* [P7]

When considered together, the archetypes lie on a spectrum where the Fully Informed archetype relies more on individual responsibility and capacity to make informed decisions, while the Trust Seeker also considers the context of organizational reputation and trust in making their decisions. In addition, the Record Keeper could be seen as an individual who wants to manage their consent decision over time, while those who do not want to review or revise their consent accept a one-time decision without records.

In addition to finding patterns based on common goals, some individuals stood out for their unique consent desires, including using more technical jargon (2/24, 8%). The use of jargon seemed to enable the process more time saving for some participants:

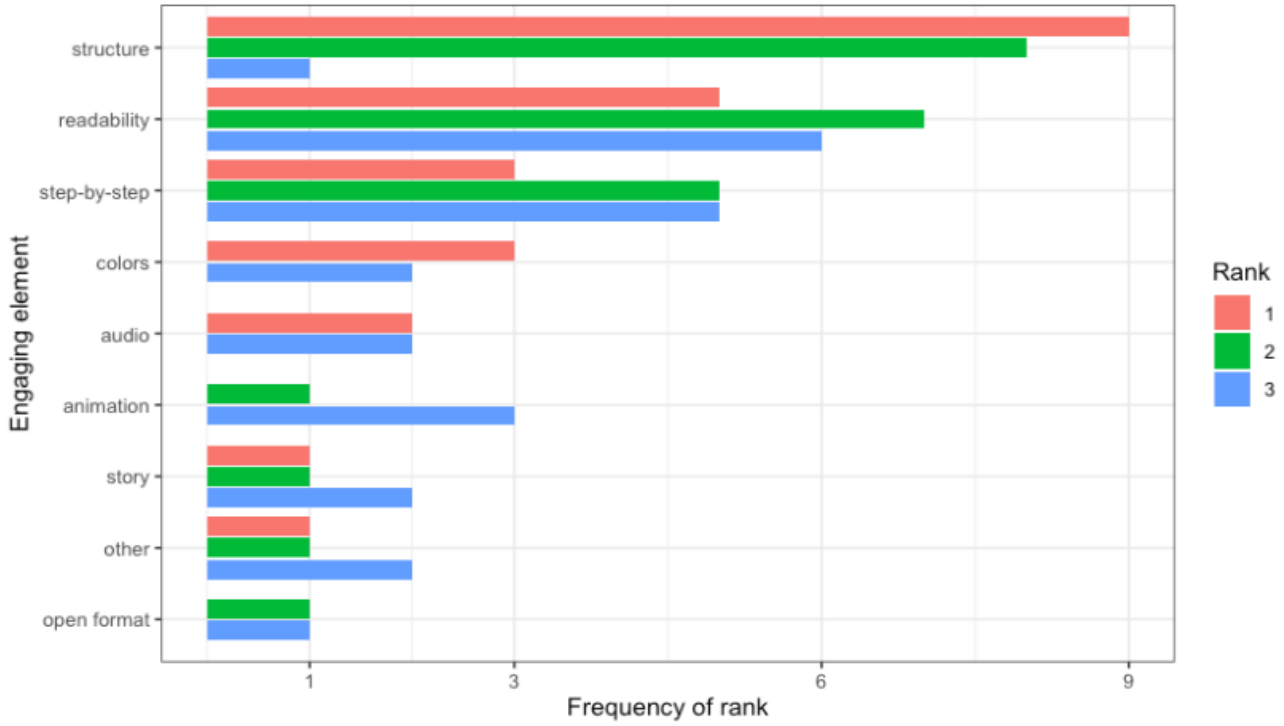
> *If I had to choose between short technical language and simple but longer language that is easy for everyone to understand, I would choose the short technical language.* [P15]

Top Engaging Elements

To better understand RQ2, about how different elements across mediums were perceived by participants, they were asked to rank the listed elements after experiencing all mediums. The most frequent element ranked first was structure, followed by readability, colors, step-by-step elements (tied with "colors"), audio, story, and others (also tied with "story"; Figure 6). The top element at rank 2 was also structure, and the top element at rank 3 was readability. When the option "others" was chosen, not all participants elaborated on what "other" element they referred to, but when they did, personal engagement (4/24, 17%) was most commonly cited. In ranks 2 and 3, structure, readability, and step-by-step were also frequently cited engaging elements.

XSL•FO
RenderX

Figure 6. The frequency of each engaging element ranked from 1 to 3 by participants.

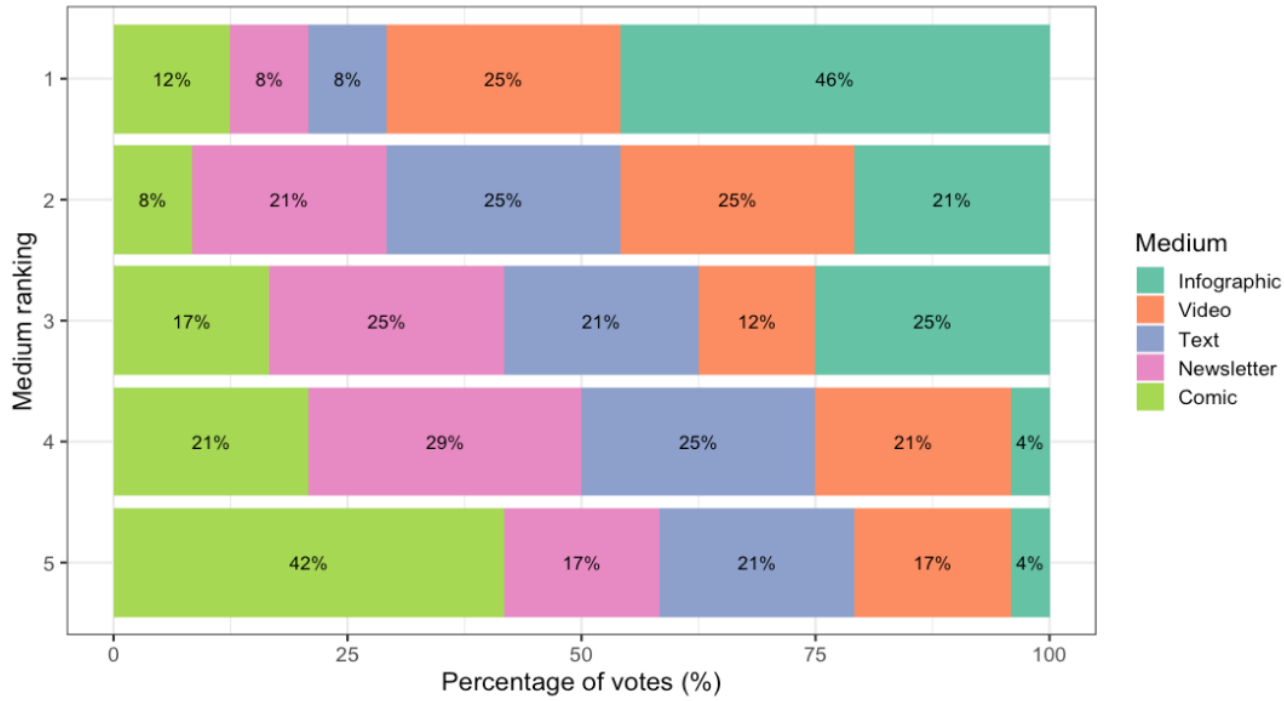


Preferred Mediums and Document Criteria

Overview

To answer RQ3, we report the results about participants' ranking for their preferred consent form after being shown each medium in Figure 7. The infographic was the overall winner and the comic the overall loser, while the video, text, and newsletter had varying trends (eg, the text was uniformly distributed across ranks 2 to 5, while the video was most often ranked 1, 2, or 4). Interestingly, no medium had consensus across the 24 participants.

Figure 7. Participant ranking of mediums (where 1 corresponds to the first choice and 5 to the last choice) by percentage of votes.



In the following sections, each medium is discussed based on (1) the top 3 factors that influenced the ranking and (2) the top 3 positive or negative document criteria adapted from good document criteria by Waller [26] (Table 1). A participant could share multiple influencing factors or document criteria. We instead looked at the number of unique coded segments within their answer. Participants could share as few (though they were prompted to try to give at least 1) or as many factors as they

XSL•FO
RenderX

desired. The coded segments for influencing factors, such as the element of time, could be positive (time-saving) or negative (time-wasting). This was to help us identify the categories that were most important to participants. Then, we contextualize the data and report whether important factors were positive or negative and their respective coded segments volume in the detailed section. Finally, Table 3 offers a summary of rank, top influencing factors, and document criteria per medium.

Table 3. Overview of the top 3 influencing factors and document criteria per medium with overall participant ranking.

	Medium				
	Infographic (rank 1)	Video (rank 2)	Text (rank 3)	Newsletter (rank 4)	Comic (rank 5)
Influencing factors					
Positive element	Understanding, time, and interest	Understanding and effort	Understanding and time-saving	Prioritization and understanding	Understanding and interest
Negative element	N/A[a]	N/A	Boring	Association with advertisements	Inappropriate fit for the context
Mixed	N/A	Time-saving and time-wasting	N/A	N/A	N/A
Document criteria					
Positive element	Numbered lists, icons, bold headings, and graphic elements	Audio, step-by-step elements, and interplay of text and graphics	Structured layout	Bolded key text, sections, and open format for skimming	Text and graphics
Negative element	Extraneous or leading icons	N/A	Lacked highlighting	Advertisement impression	Tone and audience fit

[a]N/A: not applicable.

Infographic Medium

The infographic was strongly preferred, with one-third of the participants (15/49, 31%) citing understanding as a positive influencing factor, while time and interest were in a close second. Elements such as numbered bullet points, bold headings, and icons were referenced:

> *With the bullets, you know right away what each is about in the text written underneath. In general, this is easy to grasp...* [P3]

The top 3 influencing factors were overall positive, in contrast to the other mediums.

Most positive document criteria concern design criteria such as step-by-step elements, icons, bold headings, bullet points, and color. There were much fewer negatively received elements, which were also related to the design criteria: the overuse of color and icons and the large size of the infographic. Participants had specific reactions to different icons, such as the hospital or medical professionals at the top and bottom that did not support any text or specific icons that might seem manipulative:

> *[W]ith the consent form, the "thumbs up" graphic makes it look like I'm being preempted from making a decision.* [P14]

Video Medium

Approximately one-third of the participants (14/44, 32%) reported that it influenced their understanding, followed by time and effort. Understanding was largely positive, partially due to the format that they're "forced to watch it from beginning to end, so that you perceive the whole content" (P15).

On the other hand, time was slightly more positive than negative because while most participants felt that compared to reading, the video saved time, some felt it was inefficient compared to their reading speed, or they wanted to review material but felt rewinding would be time-wasting. Saving effort was wholly positive, with participants saying that it was more accessible, entertaining, or less attention draining while still being understanding. One minor interesting influencing factor unique to the video was a perceived feeling of trust from the audio, with 2 participants mentioning that a human voice engendered confidence in the process.

More than half of the positive feedback about the video mentioned the audio element, followed by the sequential nature and use of animation and images. Less than one-fourth of the participants (12/54, 22%) liked the content, which included the interplay between text and graphics and the story element:

> *What I like about the video is that...you see movements that show what you hear at the same time via audio.* [P3]

There were about half as many negative elements as positive ones, and most were due to the video pacing. Some wanted it faster, while others wanted it slower. Interestingly, 1 participant noted the following:

> *I have the feeling that with a video like this, people are rather uncritical of the content of the consent form. One is rather tempted to agree to something. If, for example, a button appeared after the video that allowed me to consent, I would probably consent.* [P17]

XSL•FO
RenderX

Text Medium

Approximately one-third of the participants (12/42, 29%) indicated that interest and understanding were most influenced by the text. Interest was a complex influencing factor that was slightly more negative. Those stating that it negatively influenced attention felt that the text medium was boring or lacked interest compared to other mediums. The participants who viewed it positively said that the text had a simple, clean layout allowing for quick skimming, and those who felt it was neutral felt like participant 21:

> *This is the format that I know and have simply accepted by now.* [P21]

The understanding was generally a positive influencing factor, with many saying that it was clear, concise, and short; however, some felt that it was difficult to skim or that the text was confusing or dry. Some participants also felt that it saved time by being short and concise.

The most cited positive document elements of the text were the use of clear sections, headlines, and bullet points. The positive elements were twice as common as the negative elements; the majority of both positive and negative elements also stemmed from design. Participants wanted more highlighting of key facts via colored, bold, italicized, or underlined words. Less than one-third of the participants (5/18, 28%) also cited the negative impression the document gave them:

> *[I]t is still a bit boring and trivial, so you might not read it properly if you get it as a letter home, for example.* [P5]

Newsletter Medium

More than a quarter of participants (6/22, 27%) mentioned prioritization as an influencing factor, less than a quarter (5/22, 23%) mentioned understanding, and 18% (4/22) mentioned interest. Prioritization and understanding were positive influencing factors, with participants saying that the bold words and ability to skip sections allowed them to roughly understand the contents because the bold text highlighted the important information in sentences. However, the interest factor was equally mixed, with the positive influence surrounding the bolded text and headers, while the negative influence was mainly attributed to the association with advertising spam. More than half of the participants (13/24, 54%) agreed with participant 1, who stated the following:

> *It looks like advertising, by the structure and the "headline," which is repetitive.* [P1]

Although it had positive influencing factors, the negative interest likely had a large impact on the lower ranking of this form, as participant 5 said the following:

> *[I]t looks like advertising and I don't like that.... I am rather annoyed by it. The bold as highlighting and the textual design I find good.* [P5]

The newsletter's positive elements were largely regarding the design and use of structure, headings, bold text, sectioning, and the open format for skimming. The negative elements also similarly mentioned the design criteria because it looked like advertising based on prior experiences. The use of color was also disliked because the black header was too strong and off-putting.

Comic Medium

The main influencing factor was understanding, with one-third of the participants (7/24, 29%) mentioning it both positively and negatively. A slight majority (4/7, 57%) cited a positive influence on understandability. Interest was generally a positive influencing factor because the medium was novel. Less than one-fourth of coded segments showed that the comic had an overall negative influence on skimming, as the narrative-driven step-by-step format made it difficult to prioritize, reread for specific elements, or gain a quick overview. In addition, many participants disliked the comic medium as a whole, even if they could find some helpful design elements:

> *I found the comic a bit inappropriate for the topic...the message is better visualized by the little pictures, which may be better remembered but I don't like it.* [P20]

Almost half of the positive feedback for the comic stemmed from the support of text with graphics, narrative elements, and illustrations. A third felt that the tone and audience fit suited them. However, negative impressions were almost double the positive ones because audience fit and tone were unsatisfactory for more than half of the participants (14/24, 58%):

> *I'm out of the age where I still like comics.... I don't feel like I'm being taken seriously as a customer with a consent form like this.* [P16]

Participants suggested that children or older adults might be a better audience fit. Other negative feedback arose from the impression and graphic elements concerning the execution of illustrations, legibility, and lack of structure.

Discussion

Principal Findings

Qualitative analysis of participant desires for health consent revealed 3 archetypes: the Fully Informed, the Record Keeper, and the Trust Seeker. All participants wanted a high level of understanding before the consent decision, with some valuing additional elements such as obtaining copies of their decisions for their records and the trustworthiness of institutions like hospitals. The participants greatly stressed the need for short, concise, and direct consent forms that should not be longer than a page. Our results support the results of other authors, who have found that participants want to skim consent forms because consent documents are all the same, they want to save time, or they trust the ethical review of the related study [40]. In other words, individuals often engage in a form of strategic reading [26] instead of relying on attentive reading. This is why consent should contain elements that allow the visual prioritization of certain content over others, like headings, bullet points, and highlights (ie, "surface-level cues") [26] that allow individuals to skim the document effectively and discern the most important information at first sight.

On the basis of the ranking of engaging elements, the participants preferred step-by-step documents (eg, linearly

XSL•FO
RenderX

numbered lists with clear headings) instead of open or story-based formats. Structure, readability, and step-by-step elements were the top 3 engaging elements and could be easily integrated into most mediums. While our study only designed the infographic using 4 of the top engaging elements (ie, structure, readability, color, and step-by-step element), other mediums like text could also use color and step-by-step elements instead of the open-format element. However, the tone and audience fit of mediums greatly influenced participant rankings, even if some mediums enhanced understanding or visual interest (eg, comics and newsletters). The negative connotations of the newsletter with marketing and comics with childishness contributed to their low rankings, while text was seen as routine and acceptable, if boring. Instead of prioritizing one medium over another, there could be a greater focus on including the most important engaging elements within mediums (eg, adding step-by-step elements in all possible mediums).

Implications for Practice

First, the creation of data-informed archetypes can be used for better understanding, and therefore accommodating, the diverse needs of a population. To leverage information describing the use of one's own personal data as a self-determination instrument, individuals can receive contextualized information and concrete examples that are relevant to their specific needs (eg, the Fully Informed and Trust Seeker archetypes), rather than one-size-fits-all terms. Archetypes also support general audience tailoring for different goals. Different approaches to consent notices may reflect strategies to cope with the 2-fold reality stemming from the fact that the risks of consent decisions are individual, while the data sharing and processing are networked across the individual, responsible institutions, and beyond [41]. For example, the Fully Informed archetype may be more concerned with individual responsibility and the personal data processing, while the Trust Seeker wants information about the organizations involved and their security and privacy measures. Using archetypes to base user profiles could also be a way to customize their experience in meaningful macrocategories without needing to customize every possibility for individual preferences. However, more research is needed to balance the actual benefit of tailoring information to different learning styles [42] against the increased costs of its creation and implementation.

Second, different mediums can be targeted based on needed affordances (Table 3) and layered to reinforce the understanding of complex information, for example, through a multifold presentation of the same content through text, video, and infographics. Official guidance about transparency requirements' implementation [15] portrays layering techniques as an appropriate means to achieve the requirement of full disclosure while allowing for prioritization and brevity. For example, summaries containing an overview of the main clauses can accompany the more comprehensive version and can be more easily browsed while consenting, with short videos and privacy icons constituting the first layer of a written notice [20]. Distributing information on separate mediums can additionally contribute to presenting the relevant information at an appropriate time. For instance, the first layer with essential information can be displayed at the moment of the consent decision, while detailed information can always remain accessible on request [43]. However, as more guidelines for multimedia consent design arise [14,20], testing and co-designing with the intended audience is key; otherwise, a medium with a negative audience fit for certain contexts may be less effective than plain text consent (eg, the unsuitable comic medium for German adults). This can be important to test for among the intended audience, especially as comics have been a case study for cultural stigmas [44]. While they have been suitable for Indigenous populations [15], some researchers are pushing for more serious comics (similar to serious games for education) [45] and the comic co-design process itself as a research practice [46].

In terms of implementation, layering has been integrated with dynamic consent platforms. Dynamic consent was built to leverage the benefits of digital communication for health research by using digital platforms to connect people and researchers and allow participants to view, update, and change their data sharing permissions dynamically. Australia's CTRL [36], a dynamic consent platform based on open-source code, incorporates multimedia (video, illustrated text, and infographics); personalization options; and informational layering techniques. Building upon this, the layering could incorporate archetypes of general profiles to be tailored for different goals. Users of different ages may prefer different mediums, such as comics for younger audiences and videos for older audiences; similarly, users with domain expertise could choose content explained with jargon.

Finally, although we did not explicitly ask about undue influence of design elements on consent decisions and trust implications, participants clearly connected the 2, and more research is needed to better understand the deep connection. The infographic had a few complaints about specific graphics, with participant 14 saying that showing a "thumbs-up" icon was perceived as a manipulative way to preempt one from making an informed decision. Similarly, participant 17 stated that participants might believe anything shown in a video and be inclined to give consent. While guidance on ethical nudging design [47-49], as well as research on dark patterns that are to be avoided [7,8], can help shed light on such thorny issues–the issue should be more deeply studied. Considering how often human beings take decisions that are not completely rational [50,51], adding elements such as icons, color, or audio may increase the potential for manipulation of choice.

Limitations

Although we strove to obtain balanced age, education, and sex representation in our sample, they cannot be fully representative of the population. More research should be done on populations other than German adults with a larger sample size. It should also be replicated in the specific consent context of interest, as our study focused on consent to share personal data for further contact; replicating the study for clinical trial consent is important, as it may offer new archetypes, rankings, and contextual concerns. More research should also be done to study how to refine and apply archetypes in practice, as it can be insufficient or biased without continued user, expert, or patient input. Our methods only concern self-reported opinions, so there

XSL•FO
RenderX

may be a discrepancy between reported preferences and observed behaviors. The study materials may have influenced rankings and preferences, as they were generated by XD, who is not a professional designer. Therefore, certain choices (eg, the simple comic style) could have influenced participants' attitudes, making it difficult to determine the exact stimulus based on self-reported answers. While out of scope of this work, future studies can research how specific design elements or stereotypes impact rankings, for example, showing a comic with stick figures or realistic figures to German adults to better understand how to design the specific element. Before implementing consent mediums in line with applicable constraints, relevant expertise should be included in the design and evaluation of each medium.

Conclusions

To better understand the diversity of participant preferences, opinions, and emotions for IC in a health care scenario and the relevance of specific document criteria for engagement with various mediums (ie, infographic, video, text, newsletter, and comic), this study interviewed 24 individuals. The results not only have informed the generation of archetypes based on desired document features and goals but can also help create standardized consent documents that use layering to help address varying needs identified via archetypes. We also proposed recommendations for designing multimedia consent forms with a structure that promotes prioritization, such as headers, bullet points, and bold type within a contextually appropriate medium, such as an infographic or a video, so that the forms are seen by our participants as more attention-grabbing and serious than comics. It would be important to replicate this study setting in other countries, and the results could lead to contextually designed consents that align with the GDPR and other EU regulations. The findings reported here are meant to encourage further research to determine how to better involve individuals in designing useful, engaging consent forms to facilitate informed decisions concerning data sharing.

Acknowledgments

XD and MB contribution to this research work was funded by the European Union's Horizon 2020 Innovative Training Networks, Legality Attentive Data Scientists under grant agreement ID 956562. AS's contribution to this research work was partially funded by the German Federal Ministry of Education and Research and the Hessian Ministry of Higher Education, Research, Science, and the Arts with the joint support of the National Research Center for Applied Cybersecurity, ATHENE. AS's contribution to this research work was conducted partially during a research visit at the University of Luxembourg in 2021. AR's contribution to this research was financially supported by the project DECEPTICON Deceptive Patterns Online, funded by the Luxembourg National Research Fund under grant agreement IS/14717072 while she was employed at the University of Luxembourg. As an employee of the Sant'Anna School of Advanced Studies, her contribution was supported by the Biorobotics Research and Innovation Engineering Facilities" project (project identification code IR0000036) funded under the National Recovery and Resilience Plan, mission 4, component 2, investment 3.1 of the Italian Ministry of University and Research funded by the European Union—NextGenerationEU.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Interview guidelines in English.
[DOCX File , 130 KB - humanfactors_v11i1e53113_app1.docx]

Multimedia Appendix 2
Full video used in the study.
[MP4 File (MP4 Video), 18842 KB - humanfactors_v11i1e53113_app2.mp4]

Multimedia Appendix 3
English version of the infographic used.
[PDF File (Adobe PDF File), 1770 KB - humanfactors_v11i1e53113_app3.pdf]

Multimedia Appendix 4
English translation of the comic shown in the study.
[PDF File (Adobe PDF File), 2043 KB - humanfactors_v11i1e53113_app4.pdf]

Multimedia Appendix 5
English version of the newsletter shown in the study.
[PDF File (Adobe PDF File), 990 KB - humanfactors_v11i1e53113_app5.pdf]

XSL•FO
RenderX

Multimedia Appendix 6
Full codebook used in analysis.
[DOCX File , 31 KB - humanfactors_v11i1e53113_app6.docx]

References

1. Gray CM, Kou Y, Battles B, Hoggatt J, Toombs AL. The dark (patterns) side of UX design. In: Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems. 2018 Presented at: CHI '18; April 21-26, 2018; Montreal, QC. [doi: 10.1145/3173574.3174108]
2. Bauer JM, Bergstrøm R, Foss-Madsen R. Are you sure, you want a cookie? – The effects of choice architecture on users' decisions about sharing private online data. Comput Hum Behav 2021 Jul;120:106729. [doi: 10.1016/j.chb.2021.106729]
3. Santos C, Rossi A, Chamorro LS, Bongard-Blanchy K, Abu-Salma R. Cookie banners, what's the purpose?: analyzing cookie banner text through a legal lens. In: Proceedings of the 20th Workshop on Workshop on Privacy in the Electronic Society. 2021 Presented at: WPES '21; November 15, 2021; Virtual event. [doi: 10.1145/3463676.3485611]
4. Kampanos G, Shahandashti SF. Accept all: the landscape of cookie banners in Greece and the UK. In: Proceedings of the 36th IFIP TC 11 International Conference on Information Security and Privacy Protection. 2021 Presented at: 36th IFIP TC 11 International Conference on Information Security and Privacy Protection; June 22-24, 2021; Oslo, Norway. [doi: 10.1007/978-3-030-78120-0_14]
5. Utz C, Degeling M, Fahl S, Schaub F, Holz T. (Un)informed consent: studying GDPR consent notices in the field. In: Proceedings of the 2019 ACM SIGSAC Conference on Computer and Communications Security. 2019 Presented at: CCS '19; November 11-15, 2019; London, UK. [doi: 10.1145/3319535.3354212]
6. Matte C, Bielova N, Santos C. Do cookie banners respect my choice? : measuring legal compliance of banners from IAB Europe's transparency and consent framework. In: Proceedings of the IEEE Symposium on Security and Privacy (SP). 2020 Presented at: SP 2020; May 18-21, 2020; San Francisco, CA. [doi: 10.1109/sp40000.2020.00076]
7. Graßl P, Schraffenberger H, Zuiderveen Borgesius F, Buijzen M. Dark and bright patterns in cookie consent requests. J Digit Soc Res 2021 Feb 08;3(1):1-38. [doi: 10.33621/jdsr.v3i1.54]
8. Nouwens M, Liccardi I, Veale M, Karger DR, Kagal L. Dark patterns after the GDPR: scraping consent pop-ups and demonstrating their influence. In: Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems. 2020 Presented at: CHI '20; April 25-30, 2020; Honolulu, HI. [doi: 10.1145/3313831.3376321]
9. Montalvo W, Larson E. Participant comprehension of research for which they volunteer: a systematic review. J Nurs Scholarsh 2014 Nov 15;46(6):423-431. [doi: 10.1111/jnu.12097] [Medline: 25130209]
10. Niemiec E, Vears DF, Borry P, Howard HC. Readability of informed consent forms for whole-exome and whole-genome sequencing. J Community Genet 2018 Apr 31;9(2):143-151 [FREE Full text] [doi: 10.1007/s12687-017-0324-6] [Medline: 28856579]
11. McBride CM, Bowen D, Brody LC, Condit CM, Croyle RT, Gwinn M, et al. Future health applications of genomics: priorities for communication, behavioral, and social sciences research. Am J Prev Med 2010 May;38(5):556-565 [FREE Full text] [doi: 10.1016/j.amepre.2010.01.027] [Medline: 20409503]
12. D'Ignazio C. Creative data literacy: bridging the gap between the data-haves and data-have nots. Inf Design J 2017 Jul 5;23(1):6-18. [doi: 10.1075/idj.23.1.03dig]
13. Ferrer RA, Stanley JT, Graff K, Klein WM, Goodman N, Nelson WL, et al. The effect of emotion on visual attention to information and decision making in the context of informed consent process for clinical trials. Behav Decis Making 2016 Mar 28;29(2-3):245-253. [doi: 10.1002/bdm.1871]
14. Guidelines on transparency under regulation 2016/679 (wp260rev.01). European Commission. 2018 Aug 22. URL: https://ec.europa.eu/newsroom/article29/items/622227 [accessed 2024-04-23]
15. Steenberg WM. Visual communication as a legal-ethical tool for informed consent in genome research involving the San community of South Africa. University of South Africa. 2017. URL: https://orbilu.uni.lu/bitstream/10993/48262/1/published.pdf [accessed 2024-04-25]
16. Botes MW, Rossi A. Visualisation techniques for consent: finding common ground in comic art with indigenous populations. In: Proceedings of the IEEE European Symposium on Security and Privacy Workshops (EuroS&PW). 2021 Presented at: EuroS&PW; September 6-10, 2021; Vienna, Austria. [doi: 10.1109/eurospw54576.2021.00037]
17. Antal H, Bunnell HT, McCahan SM, Pennington C, Wysocki T, Blake KV. A cognitive approach for design of a multimedia informed consent video and website in pediatric research. J Biomed Inform 2017 Feb;66:248-258 [FREE Full text] [doi: 10.1016/j.jbi.2017.01.011] [Medline: 28109951]
18. Jimison HB, Sher PP, Appleyard R, LeVernois Y. The use of multimedia in the informed consent process. J Am Med Inform Assoc 1998 May 01;5(3):245-256 [FREE Full text] [doi: 10.1136/jamia.1998.0050245] [Medline: 9609494]
19. Doan X, Selzer A, Rossi A, Botes WM, Lenzini G. Context, prioritization, and unexpectedness: factors influencing user attitudes about infographic and comic consent. In: Proceedings of the Companion Proceedings of the Web Conference 2022. 2022 Presented at: WWW '22; April 25-29, 2022; Virtual event. [doi: 10.1145/3487553.3524632]
20. Rossi A, Ducato R, Haapio H, Passera S. When design met law: design patterns for information transparency. Droit de la consommation 2019;122-123(5):79-121.

XSL•FO
RenderX

21. European data strategy Making the EU a role model for a society empowered by data info/strategy/priorities-2019-2024/europe-fit-digital-age/european-data-strategy_en 2020. European Commission. URL: https://commission.europa.eu/strategy-and-policy/priorities-2019-2024/europe-fit-digital-age/european-data-strategy_en [accessed 2024-04-16]
22. Shaping Europe's digital future. European Commission. URL: https://digital-strategy.ec.europa.eu/en [accessed 2024-04-23]
23. Regulation (EU) 2016/679 of the European Parliament and of the council of 27 April 2016. Official Journal of the European Union. URL: https://eur-lex.europa.eu/eli/reg/2016/679/oj [accessed 2024-04-23]
24. Rossi A, Lenzini G. Transparency by design in data-informed research: a collection of information design patterns. Comput Law Secur Rev 2020 Jul;37:105402. [doi: 10.1016/j.clsr.2020.105402]
25. Potel-Seville M, Talbourdet E. Empowering children to understand and exercise their personal data rights. Aalto University. URL: https://acris.aalto.fi/ws/portalfiles/portal/78736878/9788855265751_Legal_Design.pdf#page=254 [accessed 2024-04-29]
26. Waller R. What makes a good document? The criteria we use. University of Reading. 2011 Apr. URL: https://uploads-ssl.webflow.com/5c06fb475dbf1265069aba1e/6090fb24421b454c776fd1a8_SC2CriteriaGoodDoc_v5.pdf [accessed 2024-04-23]
27. Pruitt J, Grudin J. Personas: practice and theory. In: Proceedings of the 2003 Conference on Designing for User Experiences. 2003 Presented at: DUX '03; June 6-7, 2003; San Francisco, CA. [doi: 10.1145/997078.997089]
28. Salminen J, Guan KW, Jung SG, Jansen B. Use cases for design personas: a systematic review and new frontiers. In: Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems. 2022 Presented at: CHI '22; April 29-May 5, 2022; New Orleans, LA. [doi: 10.1145/3491102.3517589]
29. Hendriks N, Truyen F, Duval E. Designing with dementia: guidelines for participatory design together with persons with dementia. In: Proceedings of the Human-Computer Interaction -- INTERACT 2013. 2013 Presented at: INTERACT 2013; September 2-6, 2013; Cape Town, South Africa. [doi: 10.1007/978-3-642-40483-2_46]
30. Nehme J, El-Khani U, Chow A, Hakky S, Ahmed AR, Purkayastha S. The use of multimedia consent programs for surgical procedures: a systematic review. Surg Innov 2013 Feb 14;20(1):13-23. [doi: 10.1177/1553350612446352] [Medline: 22589017]
31. Hoffner B, Bauer-Wu S, Hitchcock-Bryan S, Powell M, Wolanski A, Joffe S. "Entering a Clinical Trial: Is it Right for You?": a randomized study of The Clinical Trials Video and its impact on the informed consent process. Cancer 2012 Apr 01;118(7):1877-1883 [FREE Full text] [doi: 10.1002/cncr.26438] [Medline: 22009665]
32. Kraft SA, Constantine M, Magnus D, Porter KM, Lee SS, Green M, et al. A randomized study of multimedia informational aids for research on medical practices: implications for informed consent. Clin Trials 2017 Feb 23;14(1):94-102 [FREE Full text] [doi: 10.1177/1740774516669352] [Medline: 27625314]
33. Wang Z, Wang S, Farinella M, Murray-Rust D, Riche NH, Bach B. Comparing effectiveness and engagement of data comics and infographics. In: Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems. 2019 Presented at: CHI '19; May 4-9, 2019; Glasgow, Scotland. [doi: 10.1145/3290605.3300483]
34. Cook DJ, Manning DM, Holland DE, Prinsen SK, Rudzik SD, Roger VL, et al. Patient engagement and reported outcomes in surgical recovery: effectiveness of an e-health platform. J Am Coll Surg 2013 Oct;217(4):648-655. [doi: 10.1016/j.jamcollsurg.2013.05.003] [Medline: 23891066]
35. Ivanova D, Katsaounis P. Real-time dynamic tiered e-consent: a novel tool for patients' engagement and common ontology system for the management of medical data. Innov Digit Health 2021;1(2):45-49. [doi: 10.36401/IDDB-21-01]
36. Haas MA, Teare H, Prictor M, Ceregra G, Vidgen ME, Bunker D, et al. 'CTRL': an online, Dynamic Consent and participant engagement platform working towards solving the complexities of consent in genomic research. Eur J Hum Genet 2021 Apr 06;29(4):687-698 [FREE Full text] [doi: 10.1038/s41431-020-00782-w] [Medline: 33408362]
37. Simon HA. Designing organizations for an in formation-rich world. In: Greenberger M, editor. Computers, Communications, and the Public Interest. Baltimore, MD: The Johns Hopkins Press; 1971.
38. Ciampaglia GL, Flammini A, Menczer F. The production of information in the attention economy. Sci Rep 2015 May 19;5(1):9452 [FREE Full text] [doi: 10.1038/srep09452] [Medline: 25989177]
39. Street J, Duszynski K, Krawczyk S, Braunack-Mayer A. The use of citizens' juries in health policy decision-making: a systematic review. Soc Sci Med 2014 May;109:1-9 [FREE Full text] [doi: 10.1016/j.socscimed.2014.03.005] [Medline: 24657639]
40. Varnhagen CK, Gushta M, Daniels J, Peters TC, Parmar N, Law D, et al. How informed is online informed consent? Ethics Behav 2005 Apr;15(1):37-48. [doi: 10.1207/s15327019eb1501_3] [Medline: 16127857]
41. Hargittai E, Marwick A. "What can I really do?" Explaining the privacy paradox with online apathy. Int J Commun 2016;10:21 [FREE Full text]
42. Pashler H, McDaniel M, Rohrer D, Bjork R. Learning styles: concepts and evidence. Psychol Sci Public Interest 2008 Dec 01;9(3):105-119. [doi: 10.1111/j.1539-6053.2009.01038.x] [Medline: 26162104]
43. Schaub F, Balebako R, Durity AL, Cranor LF. A design space for effective privacy notices. In: Selinger E, Polonetsky J, Tene O, editors. The Cambridge Handbook of Consumer Privacy. Cambridge, MA: Cambridge University Press; 2018.
44. Lopes P. Culture and stigma: popular culture and the case of comic books. Sociol Forum 2006 Oct 17;21(3):387-414. [doi: 10.1007/s11206-006-9022-6]

XSL•FO
RenderX

45. Linek S, Huff M. Serious comics: a new approach for science communication and learning. In: Proceedings of the 12th International Technology, Education and Development Conference. 2018 Presented at: INTED 2018; March 5-7, 2018; Valencia, Spain. [doi: 10.21125/inted.2018.0755]
46. Tavares R, Alemany-Pagès M, Araújo S, Cohn N, Ramalho-Santos J, Azul AM. Comics in science and health communication: insights from mutual collaboration and framing a research practice. Int J Qual Methods 2023;22. [doi: 10.1177/16094069231183118]
47. Engelen B. Ethical criteria for health-promoting nudges: a case-by-case analysis. Am J Bioeth 2019 May 08;19(5):48-59. [doi: 10.1080/15265161.2019.1588411] [Medline: 31068115]
48. Renaud K, Zimmermann V. Ethical guidelines for nudging in information security and privacy. Int J Hum Comput Stud 2018 Dec;120:22-35. [doi: 10.1016/j.ijhcs.2018.05.011]
49. Meske C, Amojo I. Ethical guidelines for the construction of digital nudges. In: Proceedings of the 53rd Hawaii International Conference on System Sciences. 2020 Presented at: HICSS 2020; January 7-10, 2020; Maui, HI. [doi: 10.24251/hicss.2020.480]
50. Schuck PH. Rethinking informed consent. Yale Law J 1994 Jan;103(4):899-959. [doi: 10.2307/797066]
51. Heinrichs B. Myth or magic? Towards a revised theory of informed consent in medical research. J Med Philos 2019 Jan 14;44(1):33-49. [doi: 10.1093/jmp/jhy034] [Medline: 30649452]

Abbreviations

EU: European Union
GDPR: General Data Protection Regulation
IC: informed consent
RQ: research question

Edited by K Adapa; submitted 13.10.23; peer-reviewed by Z Wang, T Church; comments to author 27.01.24; revised version received 15.02.24; accepted 12.03.24; published 30.04.24.

Please cite as:
Doan X, Rossi A, Botes M, Selzer A
Comparing Attitudes Toward Different Consent Mediums: Semistructured Qualitative Study
JMIR Hum Factors 2024;11:e53113
URL: https://humanfactors.jmir.org/2024/1/e53113
doi:10.2196/53113
PMID:38687983

©Xengie Doan, Arianna Rossi, Marietjie Botes, Annika Selzer. Originally published in JMIR Human Factors (https://humanfactors.jmir.org), 30.04.2024. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on https://humanfactors.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Knowledge, Skills, and Experience With Technology in Relation to Nutritional Intake and Physical Activity Among Older Adults at Risk of Falls: Semistructured Interview Study

Julie Kikkenborg[1*], BSc; Emma Magelund[1*], BSc; Maria Silke Riise[1*], BSc; Lars Kayser[2], MD, PhD; Rikke Terp[3], MHSc, PhD

[1]Department of Clinical Physiology & Nuclear Medicine, Rigshospitalet, Copenhagen, Denmark

[2]Section of Health Service Research, Department of Public Health, University of Copenhagen, Copenhagen, Denmark

[3]Department of Internal Medicine, Herlev and Gentofte Hospital, Copenhagen University Hospital, Hellerup, Denmark

[*]these authors contributed equally

Corresponding Author:
Lars Kayser, MD, PhD
Section of Health Service Research
Department of Public Health
University of Copenhagen
Øster Farimagsgade 5
Copenhagen
Denmark
Phone: 45 28757291
Email: lk@sund.ku.dk

Abstract

Background: More than one-third of older adults (aged ≥65 y) experience falls every year. The prevalent modifiable risk factors for falling are malnutrition and physical inactivity, among others. The involvement of older adults in the prevention of falls can decrease injuries, hospitalizations, and dependency on health care professionals. In this regard, eHealth can support older adults' self-management through more physical activity and adequate food intake. eHealth must be tailored to older adults' needs and preferences so that they can reap its full benefits. Therefore, it is necessary to gain insight into the knowledge, skills, and mindset of older adults living at home who are at risk of falls regarding eHealth.

Objective: This qualitative study aims to explore older adults' use of everyday digital services and technology and how they acquire knowledge about and manage their nutritional intake and physical activity in relation to their health.

Methods: Semistructured interviews were conducted with 15 older adults (n=9, 60% women; n=6, 40% men; age range 71-87 y) who had all experienced falls or were at risk of falling. These individuals were recruited from a geriatric outpatient clinic. The interviews were analyzed using deductive content analysis based on a modification of the Readiness and Enablement Index for Health Technology framework.

Results: The qualitative data showed that the informants' social networks had a positive impact on their self-management, use of technology, and mindset toward nutritional intake and physical activity. Although the informants generally lived active lives, they all lacked knowledge about how their food intake influenced their physical health, including their risk of falling. Another finding was the large diversity in the use of technology among the informants, which was related to their mindset toward technology.

Conclusions: Older adults can use technology for everyday purposes, but some need additional introduction and support to be able to use it for managing their health. They also need to learn about the importance of proper nutritional intake and physical activity in preventing falls. Older adults need a more personalized introduction to technology, nutrition, and physical activity in their contact with health professionals.

(JMIR Hum Factors 2024;11:e52575) doi:10.2196/52575

XSL•FO
RenderX

KEYWORDS

eHealth; self-management; fall prevention; older adults; physical activity; nutritional intake; Readiness and Enablement Index for Health Technology; READHY; social support; support; management; fall; nutrition; diet; qualitative study; malnutrition; physical inactivity; injury; injuries; food; food intake; nutritional needs; outpatient clinic; social network; mobile phone

Introduction

Background

Among older adults, falls are common occurrences, with one-third of the population aged ≥65 years experiencing falls every year [1]. Falls are among the major causes of mortality and morbidity in older adults [2,3], and they contribute to social and economic costs as well as create a dependency on health care professionals [4]. A reduction in the incidence of falls would increase older adults' quality of life [5], relieve the pressure on health care professionals by reducing hospitalizations [6], and reduce health care system costs [4].

Malnutrition and physical inactivity are well-known modifiable behavioral risk factors for falls in older adults [7-9]. Older adults who are malnourished have a 45% higher risk of falling at least once [7]. Increased daily activity and moderate strength and intensity training 3 times a week can reduce the risk of falling by 30% [8]. Therefore, behavior change regarding nutritional intake and physical activity is important to prevent falls, which is also recommended in international clinical guidelines [8,10]. Behavior change in individuals is influenced by their self-management abilities [11,12].

eHealth, defined as "an emerging field at the intersection of medical informatics, public health, and business, referring to health services and information delivered or enhanced through the Internet and related technologies" [13], is often considered an effective tool for supporting self-management [14]. It may serve as an important "copilot" for older adults to increase their awareness of their nutritional needs and motivate them to engage in higher levels of physical activity [15-17]. To ensure that older adults adopt eHealth and find it useful, their level of knowledge, skills, and experience with eHealth must be addressed when they are introduced to it [18,19]. This includes the level of digital health literacy because it is a determinant considered important for technology adoption [20] and self-management [14].

Another important factor is the existence of social support from relatives and friends to successfully find and understand health information [21,22]. Support and encouragement from health care professionals are also essential for the sustainable adoption of health information and eHealth [22,23].

Older adults' perceptions and mindset may affect their engagement with health initiatives [24,25]. We have previously shown that older adults' perceptions and mindset affect their use of technology [6] and self-management [26]. Among older hospitalized patients, we found that a lack of awareness regarding meeting their nutritional needs was related to limited knowledge of their nutritional requirements and the impacts of food intake on their physical function [26]. Consequently, many older adults may not eat adequately [26].

The use of eHealth, the existence of social relations, and the capability to manage one's own condition, as well as perceptions and mindset, are all intermingled factors and important to include to obtain a well-functioning sociotechnical ecosystem that enables healthy behavior. In this study, we explore these aspects in the context of nutritional intake and physical activity in older adults living at home who are at risk of falling.

Objectives

The aim of this study is to explore older adults' use of everyday digital services and technology and how they acquire knowledge about, and manage, their nutritional intake and physical activity in relation to their health. These data may help differentiate among users and address specific gaps in relation to knowledge, skills, and mindset. The findings will provide insight into whether eHealth may be a feasible approach to enable and engage older adults living at home who are at risk of falling.

Methods

Overview

This is an explorative qualitative study based on semistructured interviews, which were analyzed using content analysis with a deductive approach. The study is part of a larger research program exploring how, through the use of eHealth, we can optimize older adults' self-management regarding nutritional intake and physical activity to prevent functional decline and reduce the risk of falling. An intervention outlining how health care professionals can assist and support older adults in self-management through the use of eHealth will be designed, developed, and tested. The findings of the study reported in this paper will inform the design of the intervention.

Participants and Setting

All informants were recruited through convenience sampling from a geriatric outpatient clinic in a university hospital in the capital region of Denmark. An exercise physiologist employed at the clinic obtained informed consent, allowing the first authors (JK, EM, and MSR) to contact and inform the informants before their inclusion in the study. The inclusion criteria were age ≥65 years, good comprehension of the Danish language, and being at risk of falling. The exclusion criterion was the inability to provide informed consent because of cognitive impairment. Before being interviewed, 16 informants were contacted by telephone by one of the first authors to inform them about the study and arrange the interviews. Of these 16 informants, 1 (6%) retracted their earlier decision to participate after the telephone call. Of the 15 informants, 12 (80%) were interviewed in February and March 2022; the remaining 3 (20%; all male informants to ensure information power by gaining varied experiences from both sexes [27]) were interviewed in February and March 2023 [27].

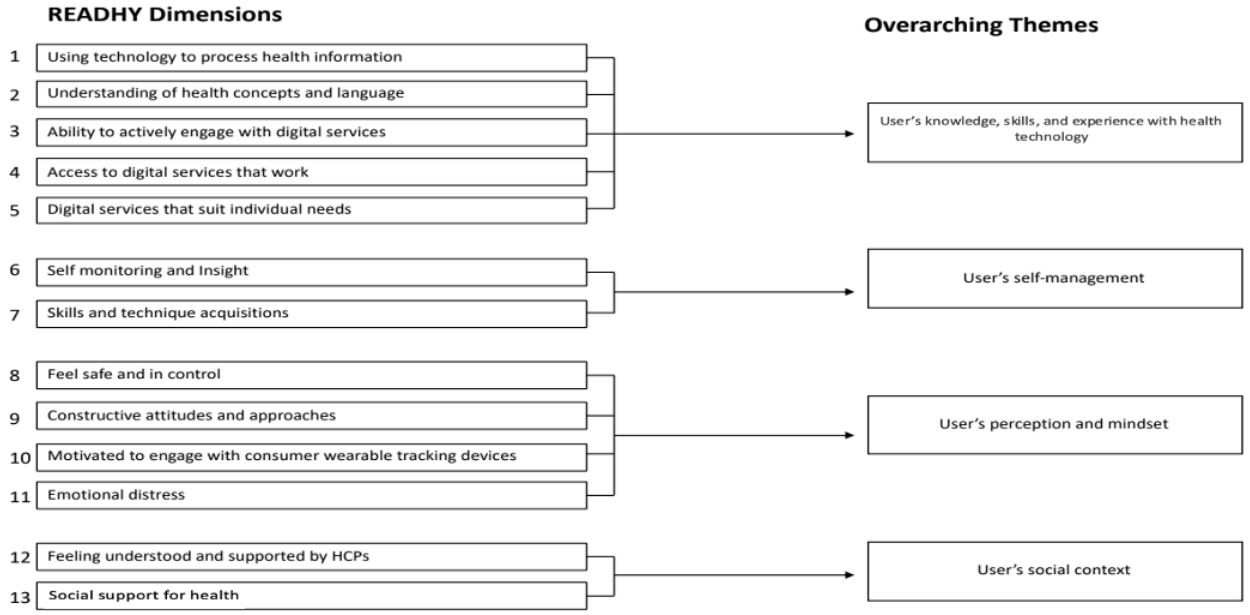
Theoretical Framework

This study is based on a sociotechnical perspective, meaning that technologies are seen as actors that interact with the users

XSL•FO
RenderX

in specific contexts instead of being passive tools [28]. To explore our informants' ability to engage with eHealth, we used the Readiness and Enablement Index for Health Technology (READHY) as the theoretical framework for the interview guide and analysis [22]. The framework can be used to describe individuals' readiness for, and enablement by, eHealth. It consists of 13 dimensions within 3 main themes: self-management, social support, and digital health literacy [18]. In this study, we applied the modified READHY framework proposed by M Blaauwhof (personal communication, November 2020). The 13 dimensions were aggregated into the following four main categories: (1) the user's knowledge, skills, and experience with eHealth; (2) the user's self-management; (3) the user's perception and mindset; and (4) the user's social context (Figure 1; M Blaauwhof, personal communication, November 2020). The main categories were used to understand how eHealth could be applied to support older adults' health behaviors regarding nutritional intake and physical activity.

Figure 1. Aggregated themes. HCP: health care professional; READHY: Readiness and Enablement Index for Health Technology.



Data Collection

Data were collected from 15 semistructured interviews using an interview guide (Multimedia Appendix 1) with open-ended questions about nutritional intake and physical activity, self-management, social context, and the use of technology. All interviews were conducted by the first authors and took place in the informants' homes. A minimum of 2 of the 3 first authors were present during the interviews. Of the 15 informants, 4 (27%) had spouses present during the interviews. The interviews lasted from 23 to 60 minutes, with an average of 42 (SD 12) minutes.

Data Analysis

The interviews were recorded using Microsoft Teams, transcribed by the first authors, and analyzed using deductive content analysis [29] based on the 4 categories from the modified READHY framework. The codes for the deductive content analysis were identified by the first authors by dividing the 4 categories from the modified READHY framework into subcategories and afterward into codes. The first authors coded the interviews using the data management software program NVivo (version 14; Lumivero). The codes were revised by the first authors and the second author (LK) to ensure reliability and to provide the option of adding newly identified codes when appropriate. The codebook consisted of 25 unique codes with a matching description (Multimedia Appendix 2). First authors JK, EM, and MSR have BSc degrees in health informatics and have experience in using qualitative and quantitative methods during their studies. JK, EM, and MSR were supervised by LK and RT, who have experience in using qualitative methods. ChatGPT (GPT-3.5; OpenAI) and Grammarly (Grammarly Inc), which offers artificial intelligence–powered writing assistance, were used to support the translation of selected anonymous quotes, the interview guide, and the codebook to improve readability and language.

Ethical Considerations

Verbal and written information about the study was provided to all informants, and written consent was obtained from all of them. The study was conducted in accordance with the Helsinki Declaration. According to Danish regulations, health science questionnaire surveys and interview studies that do not involve human biological material (section 14(2) of the Danish Act on Committees) do not require reporting to, or approval from, the Danish National Centre for Ethics [30]. All data are stored in accordance with Danish legislation (General Data Protection Regulation). The informants were not reimbursed for their participation.

Results

Informant Characteristics

A total of 15 informants (n=9, 60% women; n=6, 40% men) were included. Their mean age was 80 (range 71-87, SD 5.3) years. Of the 15 informants, 12 (80%) had experienced falls

XSL•FO
RenderX

within the last year, and 3 (20%) had experienced balance issues and dizziness. Other characteristics of the informants are summarized in Table 1, including sex, age, cohabitating status, and highest educational level attained.

Table 1. Informants' characteristics.

ID	Sex	Age (y)	Living alone or cohabiting	Highest level of education attained[a]
#1	Female	81	Cohabiting	Medium
#2	Male	72	Cohabiting	Short
#3	Female	85	Living alone	Medium
#4	Female	80	Living alone	Medium
#5	Female	86	Living alone	Short
#6	Female	80	Cohabiting	Short
#7	Male	83	Cohabiting	Medium
#8	Female	87	Cohabiting	Comprehensive school
#9	Female	79	Living alone	Medium
#10	Female	71	Living alone	Medium
#11	Female	87	Living alone	Short
#12	Male	84	Living alone	Medium
#13	Male	74	Cohabiting	Long
#14	Male	75	Cohabiting	Short
#15	Male	80	Cohabiting	Short

[a]The education variable is aggregated from the 8 levels of the International Standard Classification of Education 2011 [31] and classified into 4 categories as follows: comprehensive school (typically up to lower secondary education), short education (including upper secondary and some postsecondary programs), medium education (encompassing bachelor's and master's degrees), and long education (referring to doctoral studies).

The findings are presented in four categories: (1) the user's knowledge, skills, and experience with eHealth; (2) the user's self-management; (3) the user's social context; and (4) the user's perception and mindset.

The User's Knowledge, Skills, and Experience With eHealth

There was a large diversity in how experienced the informants were with technology and how often they used it in their everyday lives. The informants were divided into three groups based on their knowledge, skills, and experience with technology: (1) those who used technology daily (experienced informants); (2) those who used technology to some extent (partially experienced); and (3) those who had limited use of, and skills with, technology (inexperienced informants). More than half of the informants (8/15, 53%) were experienced users of technology, and 5 (83%) of the 6 male informants were experienced users of technology. The experienced informants were frequent internet searchers and daily social media users. The partially experienced informants only used technology for necessary purposes, such as checking email. Finally, the inexperienced informants had, to some extent, given up on technology either because they had no interest in using it or because they found it too difficult to use; these were also the oldest of the informants.

The experienced informants were aware of the fluctuating credibility of web pages on the internet:

> *It must be kind of random what you can do, what you find, and how you're loaded with the information you're seeking. So, you can risk finding something, I wouldn't say a lie, but something that might not be what you were searching for.* [Informant #2, male, aged 72 years]

All informants had mobile phones, and most of them (14/15, 93%) also had computers. The informants, usually the experienced ones, had smartphones (11/15, 73%), while some had nonsmartphones (4/15, 27%). The evolution of technology was unwelcome to some of the informants, and they were more comfortable using familiar technologies:

> *I don't have a modern cell phone, nor do I want one. I have this old Nokia, and it can send a text [message] and tell me what time it is; I can forward my landline calls to it whenever I'm out for a walk, so I just have it with me in my pocket.* [Informant #3, female, aged 85 years]

> *It's [the cell phone] just this one you can call from. It [the cell phone] can only do what I need it to do.* [Informant #11, female, aged 87 years]

Several informants used health technologies, such as pedometers, click-ons for hearing aids, and health information web pages. Experienced users were the ones who used technology to search for health information, but only a few used the Danish national health portal:

> *It was not long ago that I was on it [the Danish National health portal] to see all my diseases and look at my test results.* [Informant #7, male, aged 83 years]

XSL•FO
RenderX

The User's Self-Management

The data indicate that the informants in general were interested in taking care of their health, and several were making efforts to do so, including being physically active, but seemingly they paid limited attention to their nutritional intake. One informant expressed how important it was for her to be physically active for fear of losing the ability to move around freely one day. The same informant mentioned that she does not pay much attention to what she eats, arguing that she feels good now, so what she is eating must be fine:

> *Can you tell me about how you try to take care of your health?* [Interviewer]
>
> *Well, first and foremost, I do that by being physically active. And I'm so scared actually, it's something I'm afraid of, that one day I won't be able to move anymore.* [Informant #11, female, aged 87 years]
>
> *Do you ever think about what you should eat in relation to your health?* [Interviewer]
>
> *Not so much. I must admit.... I tell myself that I feel so good and I can still do so much, so what I eat cannot be completely wrong.* [Informant #11, female, aged 87 years]

Many of the informants understood and adhered to the health information they received from the internet, their relatives, health care professionals, and other sources of information, whereas only a few informants searched for health-related information themselves. Those who searched for health information typically used the internet or consulted their friends and relatives. Seemingly, the informants living with chronic conditions were more aware of seeking and appraising health information to minimize the impact of their conditions:

> *And we [friend] talked a lot about what she did...Well, but she recommended the anti-inflammatory diet, and so I did that for a longer period.* [Informant #4, female, aged 80 years]
>
> *And specifically with Parkinson's, you stiffen up, so it's especially good to stay active.* [Informant #9, female, aged 79 years]

Regarding nutrition, it also seemed that primarily those informants with diet-related conditions either sought or received information from health care professionals. Some mentioned receiving and adhering to nutritional advice from their general practitioners to better manage their conditions, such as high cholesterol, imbalanced salt concentration, and celiac disease.

Several informants set health-related goals for themselves to stay physically active in their everyday lives, such as taking daily walks, achieving a certain number of steps, and losing weight. In addition, they focused on maintaining mental well-being. Two of the informants, whose goals were taking daily walks and achieving 10,000 steps a day, expressed how they used pedometers to track their progress:

> *We walked, well, over 10,000 steps per day.* [Informant #1, female, aged 81 years]

Some of the female informants had goals related to losing weight or avoiding weight gain. Several had tried various weight reduction diets that they had heard about from relatives or the media, including commercial television programs. Diets included the ketogenic diet and intermittent fasting. In general, all informants who had health-related goals focused on achieving these goals. Some also expressed how their goal of having positive attitudes toward physical activity helped them become more physically active in their everyday lives:

> *Then I have a good [motivational] phrase: I always tell myself that what I could do yesterday, I can also do today.* [Informant #11, female, aged 87 years]

The User's Social Context

In general, the informants' relatives, particularly their children and grandchildren, were involved in, and supported them in, managing their health in terms of their nutritional intake and physical activity. Relatives inquired about the informants' health and took the initiative to help them improve it. These initiatives were often in the form of encouragement and incentives to be physically active, such as walking or using workout equipment or exercise bikes:

> *[T]hen my daughter and my son-in-law came by last Sunday and asked whether I wanted to go for a walk, and so, of course, I went with them. That happens sometimes.* [Informant #3, female, aged 85 years]
>
> *ago that I can use at home.* [Informant #12, male, aged 84 years]

The informants expressed that their relatives also played a role in shaping their eating habits, such as by introducing them to new diets or suggesting nutritional changes. However, only a few informants expressed involvement from their relatives regarding nutritional intake (compared to involvement in their physical activities):

> *[My son and daughter-in-law] have switched to a vegetarian diet 3 times a week, and we are also on board with that.* [Informant #2, male, aged 72 years]

Friends were also important sources of support. For the informants, talking with friends about their disease was a way to feel supported by discussing and sharing their experiences with peers, whereas relatives were the most important sources of support when it came to encouragement and incentives to improve their health:

> *[W]e ache and wonder how much one can actually become afflicted, as it gradually hits you. So, naturally, we discuss it a bit [with friends].* [Informant #9, female, aged 79 years]

Relatives, particularly the informants' children, were also perceived as important sources of support for technology use. They helped with the installation of new technologies and with other technical difficulties, such as pressing the wrong buttons or understanding how to use new apps:

> *[A]nd then maybe I've got a hold of something, and without knowing what it is, I fiddle and press various buttons, you know. And then it's good that I have him [his son].* [Informant #2, male, aged 72 years]

XSL•FO
RenderX

Only one informant had neither a child nor a spouse, and she experienced less social support than the other informants. This informant had no interest in physical activity, did not receive incentives from others to improve her health, and did not talk to friends about health-related topics to avoid burdening them:

> *It's not that I don't have good friends. But now, for instance, I have a very good friend who has been sick, with ataxia, and I don't want to burden her with that, you know. Because she's very sick already, so I don't want her to worry about me.* [Informant #10, female, aged 71 years]

If the informant had problems with technology, she sought help professionally, not from friends:

> *I have had some issues with my mail because YouSee and TDC [telecommunication providers] had some spam filters that hid all of my emails. I couldn't enter my email then...then they wrote to me and told me how to fix it.* [Informant #10, female, aged 71 years]

Most of the informants had positive experiences with support from health care professionals from the geriatric outpatient clinic. They felt supported by the exercise programs provided by physiotherapists from the clinic. In many cases, the positive experiences of support were related to being provided with individualized physical exercises and advice relating to their health conditions:

> *And she [the physiotherapist] has shown me some exercises to help with my balance, among other things. And I've gotten one of those round cushions that are soft, you know, so I can stand and maintain my balance.* [Informant #11, female, aged 87 years]

However, a few informants felt less supported because they had to perform a large part of the physical exercises on their own at home, and they lacked detailed instructions for these exercises:

> *Sometimes, I feel like I need an instructor.* [Informant #2, male, aged 72 years]
>
> *It [exercise instructions on paper] doesn't say; there are no measurements for the width of the board or how long you need to walk and how much you can deviate. I think it's very unspecific.* [Informant #7, male, aged 83 years]

Most of the informants listened to, and followed, the advice provided by the health care professionals from the geriatric outpatient clinic regarding their health and how to minimize the risk of falling. The advice included drinking enough fluid during the day and using training equipment at home correctly. Despite the informants' history of being at risk of falls, they generally did not report receiving information or advice from health care professionals regarding nutrition to maintain their physical function and thus minimize their risks of falling.

The User's Perception and Mindset

The informants' physical activities varied and could be classified into three groups: (1) structured physical activities, (2) incidental physical activities, and (3) inactive. Approximately half (7/15, 47%) belonged to the group that engaged in structured physical activities, such as daily physical exercises or weekly planned exercises with peers [32,33]. One-third (5/15, 33%) belonged to the group that engaged in incidental physical activities, such as house cleaning, short walks, and gardening [33,34]. Only a few (3/15, 20%) belonged to the inactive group of informants, who only did what was necessary because they seemingly had little interest in physical activities.

The informants' motivation for being physically active stemmed from the desire to remain physically mobile and avoid relying on others, as well as the opportunity for social interactions. Several informants found it enjoyable to participate in various structured physical activities together with their peers, such as going on excursions or joining training groups, rather than performing exercises alone:

> *And I'm so afraid; it's actually one of the things I'm most afraid of, that one day, I can't move, walk, or take care of myself anymore.* [Informant #11, female, aged 87 years]
>
> *And then I made a club down here, where we meet every Wednesday. And there [in the club], you're motivated to engage in a lot of events through songs, or someone comes and gives lectures. We'd go for a walk in the forest, or take trips on a bus, where we have lunch, or take some trips in which we'd have lunch out. All such events.* [Informant #8, female, aged 87 years]

A few informants from the structured physical activities group described how they did not perceive themselves as physically active. Some even expressed that they were lazy:

> *I must say that I'm living with the effects of my broken shoulder. And I'm completely lazy.* [Informant #4, female, aged 80 years]

The informants from the incidental physical activities group were familiar with how performing daily physical activities improved their health, but despite this knowledge, their motivation to perform physical activities was limited. However, there was a tendency for social interactions to be motivating factors for the informants belonging to this group:

> *Yes, and I can say that I don't think I'm fulfilling my responsibility [performing exercises] if I have to do it alone. I'm probably better at doing it with others.* [Informant #8, female, aged 87 years]

For the inactive informants, a primary reason for not exercising was the difficulty in finding personal meaning and purpose in physical activities. For some of these informants, the potential opportunity to engage in social interactions did not increase their motivation to be physically active:

> *I've always hated going on walks without a purpose. I've never played sports.* [Informant #12, male, aged 84 years]
>
> *Well, I don't like to do anything when there are a lot of people around, and when I need to stand there and do one thing after another.* [Informant #6, female, aged 80 years]

XSL•FO
RenderX

Some informants who were physically impaired because of health-related conditions found it difficult to be physically active. These informants were also either in the inactive or incidental physical activity group:

> *Then, I'll take a walk in the forest. I haven't done that very much lately, perhaps because I've had an operation on my knee. I'm not particularly fond of it, but right now, I can't go on my long walks.* [Informant #3, female, aged 85 years]

By contrast, the informants from the structured physical activities group had more knowledge about, and were more attentive to, staying active during and after an illness or injury.

Regarding the informants' perception and mindset toward their food intake, many expressed uncertainty about whether there were particular foods that would benefit their health. However, the majority expressed that they should consume more vegetables, and some also expressed that they strived to do so. Vitamins, both as supplements and in food, were highlighted as important for staying healthy. One informant expressed that protein intake was beneficial for weight loss. No informant mentioned the importance of protein intake for maintaining muscle strength and thus preventing falls. Several informants expressed acceptance regarding not necessarily adhering to recommendations for healthy eating (generally referred to as the intake of vegetables and foods low in fat and sugar) because most of them were under the impression that they ate according to their nutritional needs:

> *I actually think I eat as I should.* [Informant #3, female, aged 85 years]

Despite the informants' narratives revealing their perception of the concept of healthy food, it was evident that their food intake was largely guided by their preferences for the foods they liked. Some informants revealed positive attitudes toward allowing their preferences for tasty food to guide their intake:

> *I think about having some vegetables because it's supposed to be good for the stomach. But otherwise, I don't really think about whether it's healthy or not. I think all food is healthy if you feel like it. Unless you overdo it, of course.* [Informant #7, male, aged 83 years]

Advanced age was also mentioned as a justification for not making any changes to their eating behavior just to adhere to the consumption of "healthy food":

> *I suppose now that I've gotten so old, so...* [Informant #6, female, aged 80 years]

In general, the informants had a limited focus on their nutritional intake and on fulfilling their dietary needs. However, some female informants were more likely to report avoiding fatty food to avoid weight gain:

> *No, I don't eat fatty food, I don't eat butter on bread...It's a waste of calories because you don't need it; there's liver pâté.* [Informant #10, female, aged 71 years]

Several informants perceived technology as a necessity in their daily lives; however, some expressed that they thought technology had too much of an impact on their lives:

> *And the worst part of it is this damn computer. It starts in the morning during breakfast when I open it up, and then I'll check my email, then Ekstra Bladet [a Danish news magazine], and then Facebook, and only then am I ready for the day.* [Informant #14, male, aged 75 years]

Many of the informants had positive attitudes toward using a mobile phone because it made it easier for them to reach others and stay in contact with them. The experienced users preferred using computers or iPads because of their large screens, which made them easier to use:

> *Well, now if I should look at it [exercises on the computer] right, then I don't have to deal with the small font. It's nice with a big screen.* [Informant #1, female, aged 81 years]

Some informants expressed negative attitudes toward technology because they felt that it hindered personal communication and was challenging to manage. Several informants mentioned that negative experiences mainly occurred when the technology changed or did not work. These challenges were demotivating for the informants because it required time to resolve the problems and caused great frustration:

> *I think that, at least for us, there are too many times when it [the technology] doesn't work...and it takes too long.* [Informant #8, female, aged 87 years]

Discussion

Principal Findings

We explored older adults' use of everyday digital services and technology and how they acquire knowledge about, and manage, their nutritional intake and physical activity in relation to their health. A main finding in this study was the great diversity in the informants' experiences, mindset, and use of technology: some perceived technology as a necessity in their daily lives, whereas others viewed it as a source of frustration. Feelings of frustration were more prevalent among the oldest informants. Another main finding was that, although the informants were at risk of falls and had been referred to a geriatric outpatient clinic for the assessment and management of their fall risk, they possessed limited focus on, and knowledge about, nutritional needs to promote or maintain good physical function and thereby decrease their risk of falling. An important finding was the positive impact of the social network of relatives and health care professionals on the informants' use of technology and motivation for the self-management of nutritional intake and physical activity. There seemed to be a relationship between the informants' levels of physical activity and having a positive mindset, whereby the most active informants seemed to be more aware of the benefits of staying active and thinking positive. By contrast, there was limited focus on the importance of their food intake, which seemed to be related to the informants' limited knowledge about this topic and not to a lack of motivation to look after their health and well-being.

XSL•FO
RenderX

Comparison With Other Work

The great diversity in the informants' experiences and use of technology, with inexperienced informants belonging to the older age group, is also described in previous studies [6,35,36]. Rossen et al [35] examined readiness for technology and showed that the older age group had the lowest readiness for technology. However, the findings of Goyal et al [37] indicated that older adults may have greater adherence than younger adults once they adopt technology. This suggests that it is important to provide older adults with sufficient support to help them adopt technology. Although most of the informants (14/15, 93%) used technology, only a few (6/15, 40%) used it for health-related purposes. It was mainly the informants with chronic diseases (6/15, 40%) who searched for information on treatment and medication and generally focused on information that could minimize the impact of their chronic diseases. This is in alignment with 2 other studies and a recent review [6,23,38], which found that when older adults seek health-related information, it is often related to information about specific diseases, treatments, and medicines. Interestingly, those informants who used eHealth to find health-related information (6/15, 40%) mainly searched for information about their chronic diseases but did not search for information about how to prevent falls. This signifies that they do not consider information about preventive measures in relation to falls as being health related.

To understand individuals' motivation and ability for self-management, it is important to gain insight into their perception and mindset because these affect their motivation to engage in a specific behavior [39,40]. Our finding that older adults generally have positive attitudes toward technology is supported by the results of other studies [6,41]. The main reason for not using technology was frustration with technological challenges and because it hindered personal communication. This aligns with the results in the review by Wilson et al [19], who found that barriers to using eHealth included dislike of the technology and problems with functionality. Perceiving technological challenges may affect individuals' perception of ease of use, which is an important determinant for the intention to use technology [42]. Therefore, these potential barriers need to be addressed in future interventions. Furthermore, Wilson et al [19] found that a lack of knowledge and experience with using technology hindered use, whereas a belief in its benefits facilitated use. These findings are consistent with those of Terp et al [6], who found that the perception of technology as being useful facilitated its use among older hospitalized patients and that nonuse was mainly due to a lack of knowledge about the derived benefits from the technology. These findings suggest a need to provide older adults with knowledge about the advantages of using eHealth.

We found that the informants who were physically active (structured and incidental physical activities groups) were motivated to perform physical activities because they experienced these as being fun, they enjoyed the social element, they had been active throughout their lives, and they wished to remain physically mobile and avoid relying on others. These motivational factors for physical activity have also been described in other studies [43,44]. The reviews by Sandlund et al [43] and Bunn et al [44] found that important facilitators for commencing fall exercise programs were previous exercise habits, social support and interaction, the ability to remain independent, and the fun element. For the inactive informants, a primary reason for not exercising was the difficulty in finding personal meaning and purpose in physical activities. These findings also align with those of Sandlund et al [43] and Bunn et al [44], who found that the barriers to commencing fall exercise programs were lack of support and interest, concerns about the exercises, and unawareness of the benefits.

Our findings indicate that older adults, despite being at risk of falling, may have a limited focus on eating adequately to maintain or improve their physical health. Among several of the informants, the behavior and mindset toward food intake were focused on society's notion of an ideal slim body, which corresponds to the findings of previous studies [26,45]. Our findings suggest that this may not be due to a lack of motivation but merely due to limited knowledge; the majority were motivated and focused on maintaining their health because of their fear of losing physical function. Despite the importance of adequate nutrition, previous studies have reported limited nutritional knowledge among older adults [45,46]. Our findings indicate that older adults are more likely to receive, rather than actively seek, health-related information, a result also supported by our previous study among Danish hospitalized patients [26]. Therefore, health care professionals play an important role in providing older adults with relevant information, including information on their nutritional needs and the risk of falling. Our data revealed that older adults, in general, trusted health care professionals and adhered to the advice they provided. Our findings indicate potential benefits in ensuring that older adults receive relevant information and advice in future fall prevention interventions.

Social support from family, especially from their adult children, had a positive impact on the informants' technology use, self-management, and mindset concerning physical activity and nutritional intake. This positive influence on older adults' technology use was reported in a review by Levin-Zamir and Bertschi [21], who found that social support is paramount for many older adults in executing tasks related to health information from media sources. This finding is also corroborated by Takemoto et al [41], who found that human support increases accountability and enhances the use of technological devices. The positive impact of support from the family on the informants' mindset for nutritional intake and physical activity was also established in other studies [2,26,47]. Spiteri et al [47] reported that family support was one of the key motivators for physical activity among older adults, and Terp et al [26] found that relatives were an important resource for older adults' food intake. The positive impact of social support on older adults' self-management is also corroborated by Schnock et al [2], who found that being married or living with someone had a positive impact on engagement in fall prevention interventions and self-management among older adults. Although almost half of our informants (7/15, 47%) lived alone, our data showed that most of them (6/7, 86%) experienced social support. This is contradictory to the findings of another study [48] in which older adults who lived alone experienced less social support. Overall, in our study, relatives

XSL•FO
RenderX

were seen to have a higher impact on the informants' levels of physical activity than on their nutritional intake. This may be explained by a lack of knowledge among the relatives about the nutritional needs of older adults. In future fall prevention interventions, it is important that information about nutrition and its impact on physical function and the prevention of falls is also provided to relatives.

Several findings of this study correspond to those in our previous study, which we conducted among older hospitalized patients [6,26]. The population in the study reported in this paper differs because all informants were at risk of falling and underwent evaluation and treatment in a geriatric outpatient clinic. International guidelines [10] recommend fall prevention interventions that enhance health behavior to reduce the risk factors for falls, such as inadequate nutrition and a lack of physical activity. The informants were recruited from the outpatient clinic and had received information and training from the clinic before the interviews. We therefore expect that they were provided with information on risk factors and advice regarding optimal nutritional intake and physical activity. However, we cannot conclude from this study whether the informants had been offered such interventions. This study provides important knowledge from the perspective of older adults about their needs in terms of developing an eHealth-based intervention aimed at supporting and motivating better health behavior, thus preventing functional decline and its consequences, such as falls.

Strengths and Limitations

A strength of this study was the use of READHY as a theoretical framework because it provided a conceptual understanding of relevant aspects of individuals' readiness to engage with eHealth, such as digital health literacy, social support, and the capability to manage their own condition. The use of a modified version of the READHY framework with the addition of a fourth theme (perception and mindset) is better suited for a qualitative analysis process. Another strength is that the informants live in a country that is among the most digitalized in the world [49]. The use of digital services and technology is therefore common in the general population, and a lack of resources or difficulties in accessing the internet and eHealth are usually not barriers, as our data also indicate. Thus, as the lack of access to digital technology is not a barrier among this group of informants, it enabled us to explore how the informants acquire knowledge about their nutritional intake and physical activity in relation to their health. However, the transferability of the study is worth considering, given that the majority of the informants (11/15, 73%) lived in the same geographic area of Denmark (Nordsjælland). In future studies, the inclusion of informants from other geographic areas can help achieve greater heterogeneity.

Conclusions

This study demonstrates the potential of eHealth to support self-management in older adults, as most of them already use digital technology in their everyday lives. Older adults' age, social context, and mindset should be considered when implementing and supporting eHealth. They must be provided with knowledge about the benefits of using eHealth to improve their motivation to use it for the self-management of their nutritional intake and physical activity. Furthermore, health professionals must be aware of the need to educate older adults about the impact of nutritional intake and physical activity in fall prevention, particularly for those who lack social support.

Acknowledgments

The authors would like to thank the informants who participated in this study. The authors also thank exercise physiologist Pernille Hansen from the geriatric outpatient clinic, Herlev-Gentofte Hospital for the assistance in recruiting informants, and Marijke Blaauwhof for sharing her master's thesis work with us and allowing us to use her modified model of the Readiness and Enablement Index for Health Technology framework. The authors used the ChatGPT (GPT-3.5; OpenAI) [50] and Grammarly (Grammarly Inc), which offers artificial intelligence–powered writing assistance, to support the translation of selected anonymous quotes, the interview guide, and the codebook to improve readability and language.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Interview guide.
[DOCX File , 18 KB - humanfactors_v11i1e52575_app1.docx]

Multimedia Appendix 2
Codebook.
[DOCX File , 14 KB - humanfactors_v11i1e52575_app2.docx]

References

1. Visser M, Volkert D, Corish C, Geisler C, de Groot LC, Cruz‐Jentoft AJ, et al. Tackling the increasing problem of malnutrition in older persons: the malnutrition in the elderly (MaNuEL) knowledge hub. Nutr Bull 2017 May 08;42(2):178-186. [doi: 10.1111/nbu.12268]

XSL•FO
RenderX

2. Schnock KO, P Howard E, Dykes PC. Fall prevention self-management among older adults: a systematic review. Am J Prev Med 2019 May;56(5):747-755. [doi: 10.1016/j.amepre.2018.11.007] [Medline: 30885516]
3. Clemson L, Cumming RG, Heard R. The development of an assessment to evaluate behavioral factors associated with falling. Am J Occup Ther 2003;57(4):380-388. [doi: 10.5014/ajot.57.4.380] [Medline: 12911079]
4. Phelan EA, Ritchey K. Fall prevention in community-dwelling older adults. Ann Intern Med 2018 Dec 04;169(11):ITC81-ITC96. [doi: 10.7326/AITC201812040] [Medline: 30508457]
5. Hawley-Hague H, Tacconi C, Mellone S, Martinez E, Yang F, Su TL, et al. Using smartphone TechnolOGy to support an EffecTive home ExeRcise intervention to prevent falls amongst community-dwelling older adults: the TOGETHER feasibility RCT. Gerontology 2023;69(6):783-798 [FREE Full text] [doi: 10.1159/000528471] [Medline: 36470216]
6. Terp R, Kayser L, Lindhardt T. Older patients' competence, preferences, and attitudes toward digital technology use: explorative study. JMIR Hum Factors 2021 May 14;8(2):e27005 [FREE Full text] [doi: 10.2196/27005] [Medline: 33988512]
7. Trevisan C, Crippa A, Ek S, Welmer AK, Sergi G, Maggi S, et al. Nutritional status, body mass index, and the risk of falls in community-dwelling older adults: a systematic review and meta-analysis. J Am Med Dir Assoc 2019 May;20(5):569-82.e7. [doi: 10.1016/j.jamda.2018.10.027] [Medline: 30554987]
8. Global recommendations on physical activity for health. World Health Organization. 2010 Jan 01. URL: https://www.who.int/publications/i/item/9789241599979 [accessed 2024-04-25]
9. Gusdal AK, Johansson-Pajala RM, Arkkukangas M, Ekholm A, Zander V. Preventing falls and malnutrition among older adults in municipal residential care in Sweden: a registry study. SAGE Open Nurs 2021 Jul 07;7:23779608211026161 [FREE Full text] [doi: 10.1177/23779608211026161] [Medline: 34286078]
10. Montero-Odasso M, van der Velde N, Martin FC, Petrovic M, Tan MP, Ryg J, et al. World guidelines for falls prevention and management for older adults: a global initiative. Age Ageing 2022 Sep 02;51(9):afac205 [FREE Full text] [doi: 10.1093/ageing/afac205] [Medline: 36178003]
11. Van de Velde D, De Zutter F, Satink T, Costa U, Janquart S, Senn D, et al. Delineating the concept of self-management in chronic conditions: a concept analysis. BMJ Open 2019 Jul 16;9(7):e027775 [FREE Full text] [doi: 10.1136/bmjopen-2018-027775] [Medline: 31315862]
12. Barlow J, Wright C, Sheasby J, Turner A, Hainsworth J. Self-management approaches for people with chronic conditions: a review. Patient Educ Couns 2002;48(2):177-187. [doi: 10.1016/s0738-3991(02)00032-0] [Medline: 12401421]
13. Eysenbach G. What is e-health? J Med Internet Res 2001;3(2):E20 [FREE Full text] [doi: 10.2196/jmir.3.2.e20] [Medline: 11720962]
14. Marklund S, Tistad M, Lundell S, Östrand L, Sörlin A, Boström C, et al. Experiences and factors affecting usage of an eHealth tool for self-management among people with chronic obstructive pulmonary disease: qualitative study. J Med Internet Res 2021 Apr 30;23(4):e25672 [FREE Full text] [doi: 10.2196/25672] [Medline: 33929327]
15. Robert C, Erdt M, Lee J, Cao Y, Naharudin NB, Theng YL. Effectiveness of eHealth nutritional interventions for middle-aged and older adults: systematic review and meta-analysis. J Med Internet Res 2021 May 17;23(5):e15649 [FREE Full text] [doi: 10.2196/15649] [Medline: 33999005]
16. Muellmann S, Forberger S, Möllers T, Bröring E, Zeeb H, Pischke CR. Effectiveness of eHealth interventions for the promotion of physical activity in older adults: a systematic review. Prev Med 2018 Mar;108:93-110. [doi: 10.1016/j.ypmed.2017.12.026] [Medline: 29289643]
17. Terp R, Kayser L, Lindhardt T. An educative nutritional intervention supporting older hospital patients to eat sufficiently using eHealth: a mixed methods feasibility and pilot study. BMC Geriatr 2024 Jan 04;24(1):22 [FREE Full text] [doi: 10.1186/s12877-023-04582-x] [Medline: 38177992]
18. Kayser L, Karnoe A, Furstrand D, Batterham R, Christensen KB, Elsworth G, et al. A multidimensional tool based on the eHealth literacy framework: development and initial validity testing of the eHealth Literacy Questionnaire (eHLQ). J Med Internet Res 2018 Feb 12;20(2):e36 [FREE Full text] [doi: 10.2196/jmir.8371] [Medline: 29434011]
19. Wilson J, Heinsch M, Betts D, Booth D, Kay-Lambkin F. Barriers and facilitators to the use of e-health by older adults: a scoping review. BMC Public Health 2021 Aug 17;21(1):1556 [FREE Full text] [doi: 10.1186/s12889-021-11623-w] [Medline: 34399716]
20. Monkman H, Kushniruk AW. The consumer health information system adoption model. Stud Health Technol Inform 2015;218:26-31. [Medline: 26262522]
21. Levin-Zamir D, Bertschi I. Media health literacy, eHealth literacy, and the role of the social environment in context. Int J Environ Res Public Health 2018 Aug 03;15(8):1643 [FREE Full text] [doi: 10.3390/ijerph15081643] [Medline: 30081465]
22. Kayser L, Rossen S, Karnoe A, Elsworth G, Vibe-Petersen J, Christensen JF, et al. Development of the multidimensional readiness and enablement index for health technology (READHY) tool to measure individuals' health technology readiness: initial testing in a cancer rehabilitation setting. J Med Internet Res 2019 Feb 12;21(2):e10377 [FREE Full text] [doi: 10.2196/10377] [Medline: 30747717]
23. Brainin E, Neter E. Refined analysis of older eHealth users from an agency perspective: quantitative telephone interview study. JMIR Aging 2023 Apr 26;6:e40004 [FREE Full text] [doi: 10.2196/40004] [Medline: 37121572]

XSL•FO
RenderX

24. Archibald M, Lawless M, Ambagtsheer RC, Kitson A. Older adults' understandings and perspectives on frailty in community and residential aged care: an interpretive description. BMJ Open 2020 Mar 18;10(3):e035339 [FREE Full text] [doi: 10.1136/bmjopen-2019-035339] [Medline: 32193272]
25. Chan EY, Samsudin SA, Lim YJ. Older patients' perception of engagement in functional self-care during hospitalization: a qualitative study. Geriatr Nurs 2020;41(3):297-304. [doi: 10.1016/j.gerinurse.2019.11.009] [Medline: 31787364]
26. Terp R, Kayser L, Lindhardt T. "It is not rocket science.." - older peoples' understanding of nutrition - a qualitative study. Appetite 2021 Jan 01;156:104854. [doi: 10.1016/j.appet.2020.104854] [Medline: 32871203]
27. Malterud K, Siersma VD, Guassora AD. Sample size in qualitative interview studies: guided by information power. Qual Health Res 2016 Nov;26(13):1753-1760. [doi: 10.1177/1049732315617444] [Medline: 26613970]
28. Doolin B. Implementing e-health. In: Ferlie E, Montgomery K, Reff Pedersen A, editors. The Oxford Handbook of Health Care Management. Oxford, UK: Oxford University Press; 2016.
29. Elo S, Kyngäs H. The qualitative content analysis process. J Adv Nurs 2008 Apr;62(1):107-115. [doi: 10.1111/j.1365-2648.2007.04569.x] [Medline: 18352969]
30. Overview of mandatory reporting. Danish National Center for Ethics. URL: https://researchethics.dk/information-for-researchers/overview-of-mandatory-reporting [accessed 2024-01-25]
31. OECD, Eurostat, UNESCO Institute for Statistics. ISCED 2011 Operational Manual: Guidelines for Classifying National Education Programmes and Related Qualifications. Paris, France: OECD Publishing; 2015.
32. Caspersen CJ, Powell KE, Christenson GM. Physical activity, exercise, and physical fitness: definitions and distinctions for health-related research. Public Health Rep 1985;100(2):126-131 [FREE Full text] [Medline: 3920711]
33. Strath SJ, Kaminsky LA, Ainsworth BE, Ekelund U, Freedson PS, Gary RA, et al. Guide to the assessment of physical activity: clinical and research applications: a scientific statement from the American heart association. Circulation 2013 Nov 12;128(20):2259-2279. [doi: 10.1161/01.cir.0000435708.67487.da] [Medline: 24126387]
34. Stamatakis E, Johnson NA, Powell L, Hamer M, Rangul V, Holtermann A. Short and sporadic bouts in the 2018 US physical activity guidelines: is high-intensity incidental physical activity the new HIIT? Br J Sports Med 2019 Sep;53(18):1137-1139 [FREE Full text] [doi: 10.1136/bjsports-2018-100397] [Medline: 30786998]
35. Rossen S, Kayser L, Vibe-Petersen J, Ried-Larsen M, Christensen JF. Technology in exercise-based cancer rehabilitation: a cross-sectional study of receptiveness and readiness for e-Health utilization in Danish cancer rehabilitation. Acta Oncol 2019 May;58(5):610-618. [doi: 10.1080/0284186X.2018.1562213] [Medline: 30698060]
36. Reiners F, Sturm J, Bouw LJ, Wouters EJ. Sociodemographic factors influencing the use of eHealth in people with chronic diseases. Int J Environ Res Public Health 2019 Feb 21;16(4):645 [FREE Full text] [doi: 10.3390/ijerph16040645] [Medline: 30795623]
37. Goyal S, Morita PP, Picton P, Seto E, Zbib A, Cafazzo JA. Uptake of a consumer-focused mHealth application for the assessment and prevention of heart disease: the Under 30 Days Study. JMIR Mhealth Uhealth 2016 Mar 24;4(1):e32 [FREE Full text] [doi: 10.2196/mhealth.4730] [Medline: 27012937]
38. Zhao YC, Zhao M, Song S. Online health information seeking behaviors among older adults: systematic scoping review. J Med Internet Res 2022 Feb 16;24(2):e34790 [FREE Full text] [doi: 10.2196/34790] [Medline: 35171099]
39. Tzeng HM, Okpalauwaekwe U, Lyons EJ. Barriers and facilitators to older adults participating in fall-prevention strategies after transitioning home from acute hospitalization: a scoping review. Clin Interv Aging 2020 Jun 25;15:971-989 [FREE Full text] [doi: 10.2147/CIA.S256599] [Medline: 32612356]
40. Vincenzo JL, Patton SK, Lefler LL, McElfish PA, Wei J, Curran GM. Older adults' perceptions and recommendations regarding a falls prevention self-management plan template based on the health belief model: a mixed-methods study. Int J Environ Res Public Health 2022 Feb 09;19(4):1938 [FREE Full text] [doi: 10.3390/ijerph19041938] [Medline: 35206125]
41. Takemoto M, Manini TM, Rosenberg DE, Lazar A, Zlatar ZZ, Das SK, et al. Diet and activity assessments and interventions using technology in older adults. Am J Prev Med 2018 Oct;55(4):e105-e115 [FREE Full text] [doi: 10.1016/j.amepre.2018.06.005] [Medline: 30241621]
42. Venkatesh V, Bala H. Technology acceptance model 3 and a research agenda on interventions. Decis Sci 2008 May 09;39(2):273-315. [doi: 10.1111/j.1540-5915.2008.00192.x]
43. Sandlund M, Pohl P, Ahlgren C, Skelton DA, Melander-Wikman A, Bergvall-Kåreborn B, et al. Gender perspective on older people's exercise preferences and motivators in the context of falls prevention: a qualitative study. Biomed Res Int 2018 Jul 18;2018:6865156 [FREE Full text] [doi: 10.1155/2018/6865156] [Medline: 30112416]
44. Bunn F, Dickinson A, Barnett-Page E, Mcinnes E, Horton K. A systematic review of older people's perceptions of facilitators and barriers to participation in falls-prevention interventions. Ageing Soc 2008 Nov 03;28(4):449-472. [doi: 10.1017/S0144686X07006861]
45. Avgerinou C, Bhanu C, Walters K, Croker H, Liljas A, Rea J, et al. Exploring the views and dietary practices of older people at risk of malnutrition and their carers: a qualitative study. Nutrients 2019 Jun 05;11(6):1281 [FREE Full text] [doi: 10.3390/nu11061281] [Medline: 31195731]
46. Wegener EK, Kayser L. Smart health technologies used to support physical activity and nutritional intake in fall prevention among older adults: a scoping review. Exp Gerontol 2023 Oct 01;181:112282 [FREE Full text] [doi: 10.1016/j.exger.2023.112282] [Medline: 37660762]

XSL•FO
RenderX

47. Spiteri K, Broom D, Bekhet AH, de Caro JX, Laventure B, Grafton K. Barriers and motivators of physical activity participation in middle-aged and older-adults - a systematic review. J Aging Phys Act 2019 Sep 01;27(4):929-944. [doi: 10.1123/japa.2018-0343] [Medline: 31141447]
48. Elliott S, Painter J, Hudson S. Living alone and fall risk factors in community-dwelling middle age and older adults. J Community Health 2009 Aug;34(4):301-310. [doi: 10.1007/s10900-009-9152-x] [Medline: 19333744]
49. United Nations E-Government Survey 2022: The Future of Digital Government. New York, NY: United Nations; 2022.
50. ChatGPT 3.5. OpenAI. URL: https://chat.openai.com/ [accessed 2024-04-29]

Abbreviations

READHY: Readiness and Enablement Index for Health Technology

Edited by P Santana-Mancilla; submitted 08.09.23; peer-reviewed by M Sääskilahti, J Wilson; comments to author 08.01.24; revised version received 04.03.24; accepted 15.03.24; published 08.05.24.

Please cite as:
Kikkenborg J, Magelund E, Riise MS, Kayser L, Terp R
Knowledge, Skills, and Experience With Technology in Relation to Nutritional Intake and Physical Activity Among Older Adults at Risk of Falls: Semistructured Interview Study
JMIR Hum Factors 2024;11:e52575
URL: https://humanfactors.jmir.org/2024/1/e52575
doi:10.2196/52575
PMID:38717810

©Julie Kikkenborg, Emma Magelund, Maria Silke Riise, Lars Kayser, Rikke Terp. Originally published in JMIR Human Factors (https://humanfactors.jmir.org), 08.05.2024. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on https://humanfactors.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Health Technology Access and Peer Support Among Digitally Engaged People Experiencing Homelessness: Qualitative Study

Nóra Radó[1,2*], BA, MA; Sándor Békási[1,2*], MD; Zsuzsa Győrffy[1*], PhD

[1]Institute of Behavioral Sciences, Faculty of Medicine, Semmelweis University, Budapest, Hungary

[2]Health Center, Hungarian Charity Service of the Order of Malta, Budapest, Hungary

[*]all authors contributed equally

Corresponding Author:
Nóra Radó, BA, MA
Institute of Behavioral Sciences
Faculty of Medicine
Semmelweis University
Üllői út 86
Budapest, 1085
Hungary
Phone: 36 1 210 2953
Email: rado.nora@phd.semmelweis.hu

Abstract

Background: Although the effects of digital health are receiving wide scientific attention, very little is known about the characteristics of digitally engaged people experiencing homelessness, especially in Central and Eastern Europe. Our previous research revealed a considerable level of internet use in the homeless population of Budapest, Hungary, for general purposes (350/662, 52.9%) and medical purposes (229/664, 34.6%). Moreover, a digitally engaged subgroup was identified (129/662, 19.5%).

Objective: The aim of this exploratory study was to map out the resources, attitudes, and behaviors of digitally engaged homeless individuals in relation to digital technology to set the basis for potential health policy interventions, which will enable better access to health services through strengthening of the digital components of the existing health care system.

Methods: Between August 18, 2022, and October 27, 2022, a total of 12 in-depth semistructured interviews were conducted in 4 homeless shelters in Budapest, Hungary. Upon analysis by 3 independent evaluators, 2 interviews were excluded. The interviewees were chosen based on purposive sampling with predefined inclusion criteria. Thematic analysis of the transcripts was conducted.

Results: In the thematic analysis, 4 main themes (attitude, access, usage patterns, and solutions for usage problems) emerged. Health-related technology use mostly appeared in health information–seeking behavior. Online search for prescribed medications (5 interviews), active ingredients of medications (4 interviews), medicinal herbs believed to replace certain pills (2 interviews) or foods, and natural materials (1 interview) were mentioned. Moreover, mobile health app use (3 interviews) was reported. The intention to circumvent or check on mainstream health care solutions was mainly associated with previous negative experiences in the health care system. Several gaps in the daily use of technology were identified by the interviewees; however, more than half of the interviewees (6/10) turned out to be contact points for their peers for digital problem-solving or basic digital literacy skill enhancement in the homeless shelters. Furthermore, a lack of institutional support or special programs targeting senior clients was noted.

Conclusions: Digitally engaged homeless individuals might become mediators between their peers and comprehensive digital health programs. They have the trust of their peers, can recognize and harness the benefits of digital technology, and are able to provide meaningful help in technology- and usage-related issues through experience. Digital health services have great promise in community shelters for managing and preventing health issues, and digitally engaged individuals might be important for the success of such services.

(JMIR Hum Factors 2024;11:e55415) doi:10.2196/55415

XSL•FO
RenderX

KEYWORDS

digital health; homelessness; digital technology; internet; access; health equity

Introduction

The Digital Health Paradox

By the end of 2022, the number of mobile service subscribers climbed to over 5.4 billion people globally, including 4.4 billion people who also used mobile internet, and the usage gap has narrowed markedly in the last 5 years (from 50% in 2017 to 41% in 2022 on average) [1]. As of October 2023, around 5.3 billion people use the internet worldwide, which is equivalent to 65.7% of the total population of the world, and in the last year, 189 million new members joined the global community of internet users [2]. These are unprecedented numbers. Digitization, especially the adoption of digital health technologies at scale, has been boosted by the COVID pandemic since 2020, promising access to health care systems and beneficial health outcomes.

However, there is a growing body of evidence indicating that greater reliance on digital tools has the potential to widen the gap between those who have digital skills and access to digital tools and those who do not, thereby increasing already existing health inequalities [3]. Although digital solutions might be designed following guidelines, such as the World Health Organization (WHO) Global Strategy on Digital Health 2020-2025, which states that "digital health should be an integral part of health priorities and benefit people in a way that is ethical, safe, secure, reliable, equitable, and sustainable," certain groups are unintentionally left out of the digitization boom [4]. Paradoxically, these groups often represent patients with complex psychosocial needs, specific sociodemographic characteristics, and multiple chronic conditions, and they would benefit the most from the use of digital health technologies [5-8]. van Kessel et al [6] have referred to this as the digital health paradox.

Vulnerable Groups, Homelessness, and Health Disparities

The abovementioned groups might represent vulnerable populations that are already experiencing negative health outcomes due to their detrimental social determinants of health. This has been defined by the WHO as "the forces and systems shaping the collective conditions in which people are born, grow, work, live, and age, as well as the conditions of their daily lives" [9], and they are shaped by the distributions of money, power, and other resources [10]. Emerging research shows that there is a strong relationship between socioeconomic factors, geography, demographics, and health, with poverty, housing problems, food insecurity, abuse, gender, and ethnicity creating chronic stress, which can leave the human organism with maladaptive mechanisms that result in damage to the body's functioning systems [11,12]. These have been linked to hypertension, premature aging, cardiovascular disease, type 2 diabetes, stroke, cancer, pulmonary disease, kidney disease, and many other health problems [10,13].

In the case of people experiencing homelessness, a complex set of social determinants of health are at play, which amplify each other's impacts and leave this vulnerable group at the extreme low end of health outcomes, health care access, and health literacy. According to previous research, living without adequate housing options is associated with significantly higher rates of bacterial and viral infections, diabetes, hypertension, cardiovascular disease, mental health issues, and problematic substance use compared to populations with adequate housing options [14-16]. The COVID-19 pandemic has also increased the vulnerabilities and health risks of people experiencing homelessness [17].

Life expectancy data for people experiencing homelessness compared to the general population also support these findings. In a systematic review, Aldridge et al [18] found that socially excluded populations have an 8 times higher mortality rate for men and 12 times higher rate for women than the average population. In Western high-income countries, studies have shown that homelessness is an independent risk factor for mortality, and life expectancy varies between 50 and 65 years on average [19].

When considering health care access, homeless populations frequently experience structural barriers to obtain health care, including lack of health insurance in countries without universal health insurance, as well as competing interests in health care settings to their disadvantage alongside their own financial difficulties and competing priorities, which might lead them to secure food and accommodation before health care [17,20]. Research has also shown mistrust of health care systems and experiences of discrimination in care settings. Poorer health literacy measured among people experiencing homelessness compared to the general population might also lead to a poor self-rated health status and less adherence to medical recommendations and prescription medicines [21].

Digital Health and People Experiencing Homelessness

Previous research has shown that people with lower socioeconomic status are slower to adopt new technology, and the rates of smartphone and internet use among people experiencing homelessness were lower than the rates among those with similarly low socioeconomic status but more stable housing [22]. VonHoltz et al [23] found that while experiencing homelessness, study participants showed a 68% reduction in their likelihood to access the internet compared to when they were housed. However, in terms of preferences, it was found that low-income populations, including people experiencing homelessness, rely on smartphones rather than computers for internet access owing to cost considerations, portability, and storage issues [24]. Populations at risk for limited health literacy, as indicated in the case of the homeless populations above, are also at risk for having challenges with digital technology [25].

Previous research has mentioned that it would be beneficial to equip people experiencing homelessness with the necessary tools to get them involved in digital health ecosystems as the costs of inclusion are significantly lower than the costs of

XSL•FO
RenderX

treatment of health conditions, and the overall benefits show significance and persistence [3].

Digital Health and Homelessness: Research in Hungary

While the associations between people experiencing homelessness and their health status are well researched, especially in English-speaking countries, such as Canada, the United Kingdom, and the United States, a lot less is known about the access of people experiencing homelessness to digital health tools, their digital health literacy, their attitudes toward digital technologies, or their overall characteristics in different local settings, such as Hungary, and about the specific groups existing within homeless populations [26,27].

For these reasons, the Digital Health Research Group at Semmelweis University and the Hungarian Charity Service of the Order of Malta (HCSOM) have undertaken an overarching research agenda aiming to uncover the relations between digital health and homeless populations in Hungary. Digital health technologies are defined as "technologies which use computing platforms, connectivity, software, and sensors for health care and related uses" [5]. Previous research has mapped out the attitudes of people experiencing homelessness in Budapest, Hungary, toward telecare services, with the main finding that trust in the general health care system is the central issue when it comes to the decision of homeless populations about whether they have trust in telecare services as well [28]. This study served as a starting point for a pilot project assessing the viability of a telecare system for homeless populations [29].

Access to digital tools and digital health literacy were measured in another survey (n=662), where the results demonstrated that a significant proportion of people experiencing homelessness in Budapest, Hungary, were using the internet (52.9%), while the proportion was 81.3% in a representative sample of the Hungarian population that was used as a reference group [30]. Moreover, 69.6% of people experiencing homelessness reported mobile phone ownership, with 39.9% adding that their phone had a smartphone function and 34.6% mentioning that they have already used the internet for medical purposes [30]. In terms of self-rated digital health literacy, 24.5% rated themselves as experienced or very experienced regarding internet use, while 21.5% self-reported having mediocre experience [30].

Based on these access and skill-related characteristics, we were able to filter out a broadly defined digitally engaged group (n=129, 19.5%). This subgroup possessed their own digital tools, had some level of digital health literacy, and was partly using these digital tools for health-related reasons. When we analyzed the group and ran chi-square tests for gender, age, education, frequency of medical visits, prevalence of chronic illnesses, shelter type, and social services, the prevalence of chronic illnesses (P=.047) was found to be an associative factor in this subgroup for the likelihood of using the internet frequently for health-related reasons. However, the quantitative survey could not discern more relevant information [30].

Thus, the main aim of this study was to map out the characteristics of this specific subgroup in order to determine (1) for what purposes and (2) how the individuals in this subgroup are using digital health technologies in the framework of an exploratory qualitative analysis.

Methods

Checklist

Our methodology is based on the COREQ (Consolidated Criteria for Reporting Qualitative Research) checklist as well as the methodological framework of Győrffy et al [31] (Multimedia Appendix 1). For data collection, 12 semistructured interviews were conducted.

Ethics Approval

For all interviews, written informed consent statements were obtained, and ethics approval for the study was issued by the Scientific Research Ethics Committee of the Medical Research Council of Hungary (TUKEB 133/2020 and IV/10927/2020/EKU). In terms of the analytical framework, thematic analysis was chosen.

Recruitment

Purposive sampling was based on the following criteria: (1) presence in the social care system of the Charity Service of the Order of Malta, (2) use of the internet every second week or more frequently, (3) internet access with own smartphone, computer, or tablet or another device with a data contract, a pay-as-you-go facility, or free Wi-Fi, (4) self-rating of an average or more competent internet user, and (5) ever use of the internet for health-related reasons. The sampling criteria of this research and the filtering criteria for the broadly defined digitally engaged subgroup in our previous research matched [30]. However, the previous research involved anonymous data collection, and the present purposive sampling did not use the previous data pool as a starting point. Thus, there may or may not be an overlap between the 2 groups.

Malterud et al [32] theorized that information power can determine the ideal sample size for qualitative studies, with a sample holding more information requiring a lower number of participants. They enlisted the following 5 criteria for analyzing information power: (1) aim of the study, (2) sample specificity, (3) use of an established theory, (4) quality of dialogue, and (5) analysis strategy. In this case, the aim of the study was to assess the specific characteristics of a subgroup of people experiencing homelessness who have a digital skillset and usage pattern (see Multimedia Appendix 2 for the interview guide), thus creating a very specific sample with limited prevalence in the overall population as measured in our previous study [30]. As a result, a smaller sample size was chosen.

In the research process, 12 interviews were conducted, but in the final analysis, 10 interviews were included, which presented all the criteria of the purposive sampling specified above. Two interviews did not contain any reference to digital health usage. At this point, this might seem as a contradiction, but people experiencing homelessness may experience literacy issues, may have somewhat limited understanding due to health issues, and may have a risk of social desirability bias in relation to interview situations, which may result in self-contradictory statements,

XSL•FO
RenderX

opinions, and behaviors, in line with previous methodological findings in relation to this vulnerable population [33].

Data Collection

Interviewees were contacted by social workers or institutional assistants at 4 shelters in the social care system of HCSOM or partner institutions. These shelters either served as a night shelter (n=1) or provided accommodation on a 24/7 basis (n=3) in Budapest, Hungary.

Based on the recommendations of the social workers or institutional assistants, one-on-one semistructured interviews were conducted between August 18 and October 27, 2022.

The interview guide was developed from experiences of the previous research, the specific study aims, and a literature review. The interviews were conducted in Hungarian with a trained interviewer. The interview guide was checked on a smaller sample of the specific subgroup (n=2) and modified based on their initial feedback.

The interview guide was based on the following topics: access to and attitude toward the health care system in general, access to and attitude toward digital tools in general and usage patterns of the internet and digital tools, and access to and attitude toward digital health and usage patterns of the internet and digital tools for health-related reasons (see Multimedia Appendix 2 for the complete interview guide).

Interviews were audio recorded in person, with an average interview length of 30 minutes. All audio-recorded interviews were transcribed verbatim, and each transcript was anonymized and assigned a unique code. The interviewer checked the transcriptions for accuracy. They were not sent back to the interviewees because people experiencing homelessness struggle with literacy challenges and Thomas et al [34] argued that evidence does not support the idea that member checking increases the credibility or trustworthiness of qualitative data [34].

Analysis

Thematic analysis as described by Braun and Clarke was chosen as an analytical and theoretical framework [35]. In coding, we followed the “theoretical” technique in an essentialist or realist method, driven by the analytic interest to report about the experiences and realities of the study participants in relation to their engagement in a digital health ecosystem. In coding, we followed the deductive technique, that is, we worked with predetermined assumptions and themes, which followed the interview guide; however, clearly characterizable subthemes emerged around the previously identified main themes. Three independent researchers (ZG, SB, and NR) read and analyzed the data and discussed their findings.

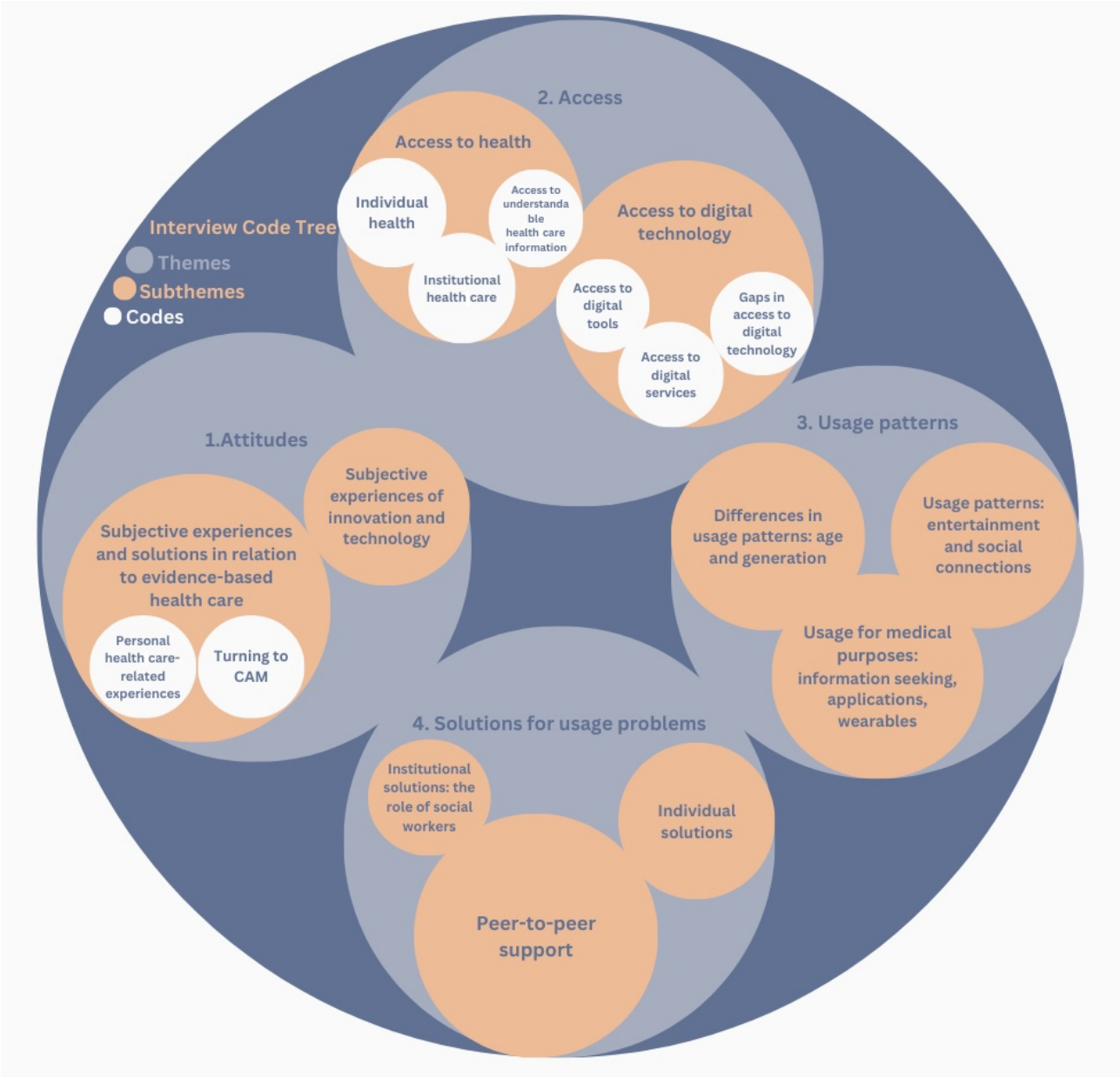
A theoretical thematic approach was used to analyze the data and identify patterns of themes based on the checklist elaborated by Braun and Clarke [35]: (1) familiarizing with the content of the data, taking notes, and making ideas for coding based on previous assumptions and following the interview guide, (2) generating initial codes manually, (3) identifying and indexing different codes across the data set manually, (4) creating relationships between the themes and subthemes, (5) defining, mapping, and naming themes, and (6) interpreting the results.

The 3 researchers discussed and developed all themes and subthemes and clarified any discrepancies during the coding. Afterwards, they laid out the final thematic map in mutual agreement. The results are supported by participants’ anonymized quotes. Interview IDs are provided for all quotes. For each interview ID, the letter indicates the first letter of the shelter where the interview was conducted (M, Miklós utca; F, Feszty; B, Budaörs; R, REVIP) and the number indicates the serial number of the interview.

For an overview of the themes, see Figure 1.

XSL•FO
RenderX

Figure 1. Interview code tree. CAM: complementary and alternative medical solutions.



Results

Demographic Characteristics

General demographic characteristics of the sample are presented in Table 1. In terms of gender, 6 male and 4 female participants were interviewed. Older age groups were overrepresented in the sample: 1 person was <40 years old, 4 people were 40-49 years old, 2 people were 50-59 years old, and 3 people were ≥60 years old. In terms of education, high school (4 people) and vocational school (3 people) were overrepresented, while 1 person had a university education and 2 people completed primary school or below.

XSL•FO
RenderX

Table 1. Demographic composition of the sample.

Characteristic	Value (N=10), n (%)
Gender	
Male	6 (60)
Female	4 (40)
Age (years)	
<40	1 (10)
40-49	4 (40)
50-59	2 (20)
≥60	3 (30)
Education	
Primary school or below	2 (20)
Vocational school	3 (30)
High school	4 (40)
University	1 (10)
Shelter	
HCSOM[a] Temporary Shelter (Feszty)	2 (20)
HCSOM Integrated Shelter (Miklós utca)	4 (40)
Shelter House Foundation's Night Shelter (Budaörs)	2 (20)
REVIP Baptist Integration Center (REVIP)	2 (20)

[a]HCSOM: Hungarian Charity Service of the Order of Malta.

Theme 1: Attitudes

Subtheme 1: Subjective Experiences and Solutions in Relation to Evidence-Based Health Care

Code 1: Personal Health Care–Related Experiences

Experiences with mainstream health care systems (hospitals, doctors, nurses, pharmacists, other medical personnel, prescription medicines, and pills) were mixed. In a minority of interviews (2/10), positive experiences with regard to access to care, quality of care, and how one was treated by the medical personnel were noted. However, the majority (7/10) reported negative experiences largely due to a negative attitude, stigmatization and mistreatment coming from the medical staff, and inadequacy of care. These signaled an overall negative attitude toward the general health care system.

> *I'm completely okay to be honest, I experienced that there are differences between the hospitals, I can only say that.* [Interview M05]
>
> *(…) …they notice where they have to go and then they have a completely different stance. Also, the emergency medical doctor, who is here, or if the ambulance services come. They behave completely differently. (…) They are condescending. Okay, we'll do it later. Okay, come back later. And another one: do pack your stuff already, we are set to go. So, they (…) are not helpful.* [Interview B09]

Code 2: Turning to Complementary and Alternative Medicine

In a minority of interviews (3/10), turning to complementary and alternative medical solutions, medicinal herbs, or Chinese medicine was noted, which was considered as an equivalent alternative of traditional Western medicine. In parallel, in 2 interviews, a negative stance toward drugs and medicines (mentioned within the same textual context) was noted.

These shed light on the fact that interviewees sought out different potential solutions to their medical problems as some of them experienced that health care systems and traditionally produced drugs cannot and have not so far provided them with appropriate solutions. They had taken medicinal herbs or trusted ingredients, which were recommended by a trusted person or were found online.

> *I am aware of that, I looked up the side effects, the medicines, I will not take what they prescribe. I have already played along for long. I rather drink an herbal tea.* [Interview F01]
>
> *I can feel if something's off in my body, and then I look up certain things, but to be honest, I always start with medicinal herbs, and not with pills. I go to the pharmacy, and I look up on the internet what is recommended for example for lower abdominal pain or for a story with joints.* [Interview M07]

XSL•FO
RenderX

Subtheme 2: Subjective Experiences of Innovation and Technology

Subjective experiences and attitudes toward novelties and technology were mixed. In almost half of the interviews (4/10), openness toward trying new programs and applications appeared, while in 2 interviews, a complete lack of interest was reported.

Attitudes toward the use of digital tools and the internet were also mixed. In part of the interviews, lack of trust and negative experiences were reported, for example, the risk of data misuse (1 interview), risk of making mistakes due to the autocomplete function and the speed of digital tools (1 interview), and inaccuracy of step counting (2 interviews). In another set of 4 interviews, openness toward trying new programs and applications appeared, while in a minority of interviews (2/10), lack of interest in this area was reported.

> *As I'm homeless at the moment, I don't have enough (money) on my pay as you go facility that I could use the internet unlimited. Where there is free Wi-Fi, I certainly search for things I think of or what I gather from my environment, or from my godchildren. So, I want to keep up with today's world in spite of the fact that I'm now a little bit on the brink of it.* [Interview M07]

> *You can really misuse data. I had that now, as well. Someone tapped into my bank account, abroad. I had to block access to my debit card, and I will have it done at some point.* [Interview F01]

> *As this device (tablet) works so that if my hand starts to shake just a little bit, and it gets close to it, it pulls in. And then, it writes something that I don't want to. So, I don't think that it is so reliable.* [Interview R12]

Theme 2: Access

Subtheme 1: Access to Health

Code 1: Individual Health

The majority of interviewees (7/10) self-evaluated their health status as average or worse. Chronic diseases (cardiovascular and heart problems, and type 2 diabetes), cancerous tumors, and lasting harm from injuries were characteristic for the group. In 4 interviews, managed alcohol problems were reported. Drug abuse was not mentioned, and in 2 interviews, aversion to drugs was noted. Diagnosed mental health problems were not mentioned.

Some interviewees regularly took medicines or mentioned that their doctors prescribed them certain types of medications, which they did not take. Some interviewees made decisions in medical matters based on their own opinions and beliefs without any professional evidence.

> *My troubles look like heart, liver, kidney, arterial obstructions. I had deep vein thrombosis in both legs but I carried that for long. I have a very high tolerance for pain. I usually operated on myself. I froze both of my legs and I cut the ulcer out as deep as I could. Then I put herbs into the wound. It recovered within 2 weeks.* [Interview F01]

Code 2: Institutional Health Care

Interviewees were clients of 4 homeless shelters in Budapest, meaning that they had institutional access to basic health care. Their legal social security status could be provided by social institutions on the grounds of homelessness under Hungarian law. Accessible health care services included primary care (prescription and dispensing of medicines, referral to specialists, and care work), publicly funded specialized outpatient care, inpatient (hospital) care, and rescue in case of emergency.

> *If I have any problem, the Maltesers (HCSOM) have a doctor's office. And if I can go there on my own feet, then I go there. If you can't, then you will be transported to the hospital by default. There are decent people who help or call an ambulance. In the doctor's office, they refer you to any specialist, no matter whether it's dermatology or cardiology. Thus, they can get you to any kind of specialist.* [Interview F03]

Code 3: Access to Understandable Health Care Information

In at least one interview, a lack of access to understandable health care information was reported, and several interviewees pointed out that they were seeking out medications and ingredients online with the help of digital tools in order to understand what impacts those materials had on their body. The need for understanding health-related information was noted in at least half of the interviews in certain forms, for example, they looked up prescription medicines (5 interviews) and their ingredients (4 interviews) online, and in at least one case, they did that for their family members as well.

> *(….) most of the time, physicians use such Latin words in general, as lawyers do. Make it simple! No one is that much overeducated to know these. For example, laboratory tests. They should include what does this mean, sodium was X. There are some apps where you can look that up.* [Interview F01]

Subtheme 2: Access to Digital Technology

Code 1: Access to Digital Tools

An overwhelming majority of interviewees (7/10) used smartphones. Notebook use was reported in 1 case, and tablet use was reported in further 2 cases. One interviewee reported power bank use to charge the device.

In a minority of interviews (2/10), it was reported that in times of need, phones, tablets, and computers were sold; thus, these were not permanently accessible tools.

> *In this living situation, people get such digital devices much easier off their hands, if they are not in such a whacking need of them, simply to be able to make money out of it.* [Interview F04]

Code 2: Access to Digital Services

In homeless shelters, interviewees had access to the computers in possession of the shelters, and through those devices, they

XSL•FO
RenderX

could get access to the internet. In certain shelters, free Wi-Fi and the option to charge their phones were available.

The majority (6/10) used free Wi-Fi inside and outside of the shelters and looked actively for options of free Wi-Fi. They could afford subscription (3 interviews) or pay-as-you-go facilities (5 interviews) less frequently. In some cases, the interviewees reported that they visited cafes in order to be able to charge their phones or use the internet (3 interviews).

> *(…) the Wi-Fi is so strong that you don't have to go in and consume something, or if you go in and drink a cup of coffee or water, you get the Wi-Fi password, then sit in front of it on a bench, and it has such a strong signal that you can use it there as well, until it is open.* [Interview M07]

Code 3: Gaps in Access to Digital Technology

The interviewees reported both tool supply and network coverage as existing problems. Several interviewees mentioned the need for securing a device (smartphone) or asked about whether there was potential for decreasing the price of subscriptions and pay-as-you-go facilities. The presence of smart benches in public spaces was mentioned in 1 interview, and free Wi-Fi on trams and busses in Budapest was mentioned in another interview.

The computer park and Wi-Fi network coverage in the shelters were not mentioned as problems in the majority of interviews (7/10), and the idea of having more connectors in the building to allow easier charging surfaced in 1 interview.

> *Some support would be great so that a basic device could be ensured for them. And a separate health network, which is for free. For people who are ill. As there are these crisis helplines and these have green numbers.* [Interview F01]

> *… prices could be reduced (…) and for example such benches could be installed where phones can also be charged. And then you could use the Wi-Fi there.* [Interview F03]

> *It's very difficult, I would say there could be more charging stations. The bigger shopping malls are covered, that's fine, but what if you suddenly notice your phone is dead and you cannot go into any such places, or you are far (from the charging station), and a homeless cannot buy a ticket… How do you go there?* [Interview M05]

> *I would tell you the truth… I'm sure it would be feasible to have free Wi-Fi on busses and low-floor trams. So here, we have Wi-Fi, since this is a shelter but when we go 20 meters further, there isn't any, the network disconnects.* [Interview M05]

Theme 3: Usage Patterns

Subtheme 1: Differences in Usage Patterns: Age and Generations

Every interviewee used the internet at a measurable frequency on their own device. The age of the participants ranged from 35 to 69 years. The interviewer did not explicitly ask about usage characteristics by age, and the topic came up spontaneously in the case of 6 interviewees when talking about attitudes toward novelties.

In several cases, the interviewees mentioned generational differences in usage, characterizing the older generation as less involved in the digital world and less interested in novelties, while younger people were considered to be already born with digital devices, and their usage seemed to be self-evident. In 1 interview, it appeared that if there was individual motivation, then age would not pose a hinderance with regard to usage.

> *This is a fundamental thing, really, but many don't know, especially the older generation. (…) So, I'm quite digital, but I'm only 40 years old for that matter. We grew up on these devices more or less already.* [Interview F03]

> *(…) I think this is age-dependent, thus generation-dependent. The elderly are okay with their basic phones. When it rings, they pick it up, then put it down. My generation already needs it more, we use it more often and the younger even more, they don't even put it down.* [Interview M05]

Subtheme 2: Usage Patterns: Entertainment and Social Connections

Interviewees mainly used the internet for entertainment and maintaining their social relationships. Watching movies, listening to music, reading e-books, and playing phone-based games were also reported. Seven interviews mentioned Facebook and 1 interview also mentioned X (or formally Twitter) as frequently used social media sites. A minority of interviews mentioned information gathering through Wikipedia (1 interview), reading news (1 interview), and online banking (1 interview) as use cases.

> *I watch movies, and look up e-books, in a topic that I'm interested in. Mostly self-healing, quantum healing and such banalities.* [Interview F01]

> *I had a smartphone, so not only the music, YouTube, Facebook page is important to me, but also Wikipedia, where I can look up everything, or for example, I read a lot about various things, and the disease that I had. This is very important to me.* [Interview M05]

> *Interesting that I also keep in touch with my physician via e-mail. I had for example a CT scan, and then everything worked entirely online. I received my appointment and also the findings online. I also consider this a very positive thing, so that it is also in the cloud, and they can see it, the whole thing is much easier… I just give them my social security card (TAJ-card), and then I tell them what prescribed medication I want to have. So, I consider this absolutely positive.* [Interview M05]

Subtheme 3: Usage for Medical Purposes: Information Seeking, Applications, and Wearables

In several interviews, information seeking for medical purposes was reported. For example, interviewees looked up prescribed medications (5 interviews), active ingredients of medications

XSL•FO
RenderX

(4 interviews), medicinal herbs believed to replace certain pills (2 interviews) or foods, and natural materials (1 interview). One interviewee mentioned purchasing a product believed to have medicinal value online on the basis of a Facebook advertisement.

One interviewee in their 30s communicated with the doctor about health problems via email, provided information about their illness and the prescribed medicines online, and used a health app and a step counter. These tools (health app and step counter) were also mentioned by 2 other interviewees, but one of them stopped using the step counting option as they believed it was inaccurate.

> *I look up the active ingredient of a pill, for example when before chemotherapy certain medicines were prescribed for me, and I looked up what kind of active ingredients they have, what side effects could they have, because a package leaflet is one thing and a real person who already had this experience and took the medicine, and what is their opinion, is another thing.* [Interview M05]

> *I already had this step counting thing, this daily fitness thing. And I remember I had a heart rate monitor in my old Samsung S5, and now I really miss that my current phone doesn't have that anymore. (…) I also use a menstruation tracking app.* [Interview M05]

> *I usually look up online for my partner what kind of cremes and medicines there are … if they are interested what kind of ingredients the pill has, and due to his blood pressure.* [Interview M06]

Theme 4: Solutions for Usage Problems

Subtheme 1: Individual Solutions

We included interviewees in this study who previously stated that they frequently used digital tools and self-evaluated their skills as at least average. The majority of interviewees (8/10) themselves did not mention usage problems, and when they had problems, 1 interviewee asked their family members for help but added that they preferred to solve their problems on their own.

Subtheme 2: Peer-to-Peer Support

It was frequently (6 interviews) reported that the interviewees offered their help to other clients who lived with them in the same shelter if they had trouble around the usage of digital tools or the internet. They solved usage-related problems for their peers, such as registration of SIM cards, activation of pay-as-you-go facilities, antivirus actions for devices, problems around online programs like Facebook and Messenger, and questions around online purchases. These user troubles represented basic problems, and the majority of interviewees (8/10) had the knowledge and skills to solve them.

> *Last time they wanted to buy something online, and they asked my help in that. (…) Now one of the guys from the shelter came up to me how to activate the SIM card. And then I activated it for them. Such issues are always in need.* [Interview F04]

> *There were some who asked me how to log in, how to register with an email address, how can they make a Facebook profile. Then I helped first to make an email account and then to register with that. (…) I was happy that I could help and they accepted it gladly. And then I saw that they were using it very well, they were glued onto their screens and were happy about it.* [Interview M05]

> *Usually Facebook, Messenger, or when they cannot download a game. And there is an antivirus program on every smartphone with a broom icon but they don't know what that is. So, I tell them, pick it up and swipe with it. Clean it. And then they look at me confused. Okay, give it to me. So, then I do it, and they look. Wow, then they say, it went down to zero. Yeah, and then I say that's the point, not to have anything on it. So there are always things like this.* [Interview B09]

Subtheme 3: Institutional Solution: Role of Social Workers

Interviewees did not report institutional solutions aiming at the development of digital skills. In 1 interview, a social worker was mentioned who provided the client with basic information on tablet use. In this case, it was the individual initiative of the social worker and not an element built into the given institution's services.

> *(…) then the social worker came up to me, and taught me the basics, and then they said that I should now keep pressing the buttons around nicely, and then I'll figure everything out by myself.* [Interview R12]

Discussion

Digital Technology and People Experiencing Homelessness

Digital technologies show a general potential for improving patient outcomes. For example, Bruce et al [36] showed that both clinical and patient-centered care outcomes were significantly better with the use of mobile health technology among 2059 orthopedic patients. However, according to a systematic catalog on digital health systematic and scoping reviews, there is less specific evidence on equitable health care (16.7%) [37].

In relation to the homeless population and digital technology, Heaslip et al [26] identified in their systemic review that mobile technology has a measurable health impact on the homeless population directly and indirectly. As an indirect impact, maintaining relations with relatives and friends as well as the outside world through entertainment, movies, and music strengthened their social connectedness and elevated their self-esteem, which in turn can have a positive impact on their personal health [38]. For the direct health impact of digital technology, they found limited evidence, with the main areas being reminders for repeat prescriptions or health care appointments. However, Heaslip et al [26] mentioned that the homeless population appears to consider that digital technology has potential health benefits, mostly in terms of online health information support and appointment reminders.

XSL•FO
RenderX

Our results partly strengthen these findings. The interviewees in our digitally engaged homeless subgroup used their digital tools primarily for entertainment purposes and to maintain their personal relationships. In terms of health care, they used their devices as new channels to reach solutions for their health problems outside the conventional health care system and to search for health-related information. However, most interestingly and most importantly, the majority of interviewees (6/10) shared that this subgroup is supporting their peers in taking up digital skills and is helping them solve their usage- and device-related problems, and this behavior has a lot of untapped potential for widening digital health usage in the homeless population.

Health Care Needs and Personal Experiences

As indicated by the demographic characteristics, older and predominantly male interviewees shared their experiences. Consistent with the results from our previous studies [28-30], the majority of interviewees (6/10) reported multimorbidities [39,40] and having chronic diseases, such as cardiovascular diseases [41], type 2 diabetes, cancer, and permanent injuries. Older age (≥50 years) was associated with worse physical health in the homeless population, which was noted in the interviews, as the self-reported health status was regarded as average or worse [19].

In our small sample, there was no mention of mental health problems other than addictions. Previous research found that the ratio of serious mental disorders among people experiencing homelessness in Hungary was very high [42], which is in line with findings from Western countries [43]. Underdiagnosis and undertreatment of mental health problems caused by stigmatization and underperformance of the Hungarian care system might be prevalent among our interviewees as well [44]. Moreover, in line with previous studies, which estimated the prevalence of alcohol abuse at 8.5%-58.1% [45], treated alcohol problems were noted in 4 interviews; however, illicit drug use or treated drug abuse problems were not mentioned. A systematic review found that alcohol abuse is more prevalent in mainland Europe [43].

Issues of Access to Health Care and Digital Tools

Access to primary care is resolved via the care settings of the Health Center of the HCSOM, which includes prescribing drugs, providing basic care services, and referring clients to specialists. In line with previous studies [17,20], the experiences of interviewees with accessing health care were mixed.

When looking at access to digital tools and digital services, in line with previous research, the majority of interviewees (7/10) had smartphones, which are more accessible to people with a low socioeconomic status [24]. The partial accessibility of digital devices and their use as assets in times of need as described in a minority of interviews (2/10) have been mentioned by Heaslip et al [26]. As a need, device supply was primarily mentioned by the participants, and this is in line with our previous study where 21.4% of respondents mentioned lack of a smartphone as the main barrier for not using the internet and 24.1% mentioned that availability of an appropriate device would help them use the internet more [30].

Digital services, such as computers of the shelters, were available to the participants, and in some shelters, free Wi-Fi or charging was also provided. The majority of participants (6/10) looked for free Wi-Fi options outside the shelters as well. One interviewee mentioned the lack of free Wi-Fi on public transport services and the lack of installation of smart banks in Budapest as barriers to usage. Such infrastructural problems were mentioned as causes of nonusage by 7.6% of respondents in our previous study [30]. On the other hand, several interviewees mentioned using the paid services of cafes to charge their phones or use Wi-Fi.

Several interviewees also mentioned the need for a potential decrease in internet service prices or device prices, which is in line with the finding of our previous study where 18.4% of participants said that better access to free Wi-Fi, pay-as-you-go facilities, or data contracts would help them use the internet more [30].

Problems Around Trust

Some interviewees mentioned the feeling of being unwelcome in conventional health care settings, which is in line with previous research [41]. Some of them mentioned difficulties in getting appropriate treatment and a negative attitude from health care personnel, which might negatively influence their desire to seek health care in the future and their overall trust in the health care system, and this might explain their turn away from mainstream health care solutions.

These aspects might include a negative impact on medication adherence and an overall mistrust in mainstream medical solutions, such as taking antibiotics and chronic disease drugs, with a turn to alternative solutions. From the interviews, it was found that managing treatment themselves instead of relying on medical personnel based on their own beliefs without medical evidence was a solution. Moreover, turning to alternative and complementary medical solutions, such as homeopathy, herbal medicine, and Chinese medicine, was a way to express mistrust in conventional care settings, and digital solutions can open up a channel outside of the conventional health care system to reach such alternative solutions.

Mistrust and negative attitudes toward the health care system coupled with the need for understanding health-related language, prescription drugs, and active ingredients were associated with the main health-related use of digital tools and services in the majority of interviewees (8/10).

Age as a Predictor for Usage and Openness

When asked about usage patterns, several interviewees spontaneously shared their views on how age differences matter in usage prevalence, outlining that older generations might be less involved and less interested in novel technologies. Several studies, including our previous quantitative research, support that age is a key sociodemographic variable that has an impact on use [29,30,46,47]. Our quantitative data showed that in access to technology, age did not seem to be a key factor; however, it might be considered as a significant factor when self-evaluating competence in digital literacy skills. This appeared in at least one of the interviews, with the respondent explaining less elevated technological skills with age.

XSL•FO
RenderX

At least three interviews indicated that age was associated with openness toward or willingness to try new technologies, which might be in line with the findings of a representative questionnaire survey (n=1500) on digital health–related knowledge, attitudes, and needs [46]. This survey was completed in 2021 and found that a quarter (26.5%) of individuals aged 65-74 years and a third (31.9%) of individuals aged older than 75 years would not like to try digital technologies in the coming years [46].

Lack of Systematic Support Results in Peer Support for Skill-Related Problems

While interviewees recognized some support from shelters in solving infrastructural and service-related technology issues, there was a perceivable lack of systematic solutions when it came to usage-related problems and digital literacy issues. Only 1 interviewee mentioned that a social worker helped them set up their tablet and navigate through basic usage scenarios.

As we selected interviewees based on at least average self-reported digital health literacy skills and aptitude toward digital technology, with some demonstrating previous educational or professional background in IT services, their less digitally skilled peers turned to them for help.

The majority of interviewees (6/10) provided unintentional peer support in relation to technology usage issues, solved technology-related problems, and provided guidance for future scenarios. Peer support, also in this context, is defined in the literature as a process whereby individuals with lived experiences of a particular phenomenon provide support to others by explicitly drawing on their personal experiences [48]. Intentional peer support works as a formalized framework of this process that is fostered and developed by institutions, while unintentional peer support remains under the radar of institutions. The literature recognizes the potential of peer support and peer support workers, who have the necessary training and provide intentional support to their homeless peers by sharing their lived experiences in different areas of life, and members of this digitally engaged subgroup might show potential for offering peer support in digital upskilling [48,49]. Moreover, anyone considering a comprehensive digital health program for homeless groups in Hungary that concentrates on offering solutions to infrastructure and skill-related problems should take into account the untapped potential of members of digitally engaged subgroups. These individuals, through their elevated trust levels among peers, might provide better outcomes in digital upskilling than official and institutionalized digital health literacy programs. A systematic review found that empowerment and self-esteem in the homeless population increased when working with homeless peers as mentors and educators, and that peer support in general facilitates acceptance of illness and recovery and increases efficacy, social skills, and coping [50].

Strengths

Through the qualitative analytical framework, the characteristics of a unique subgroup of digitally engaged people experiencing homelessness could be explored in a less studied area of digital health for equitable health care, where systematic mapping of review studies showed notable gaps in evidence [37].

The study aimed to enrich the still relatively small body of research concerning the characteristics, including the digital health–related characteristics, of the homeless population in Central and Eastern Europe. In North America and Western Europe, where the majority of studies involving the homeless population are conducted, the demographic composition of such populations as well as the health care system may differ significantly from Hungarian experiences, with different problems and solutions at individual and systemic levels.

Limitations

Our study has certain limitations. As a qualitative study using in-depth semistructured interviews, the sample size was small, and this should be taken into account when drawing inferences. The study participants represented the urban homeless population from Budapest, Hungary, where socioeconomic conditions might differ from those in the countryside. The recruited homeless people had a living connection to the social infrastructure; therefore, rough sleepers and other people who were not connected to any social initiatives were not represented. The research team exclusively relied on self-reporting of digital tool access and use, and did not attempt in any way to verify these reports (eg, via phone bills, direct observation, and other methods).

In relation to people experiencing homelessness, there is an increased risk of social desirability bias when conducting interviews, meaning that respondents tend to modify their responses in the presence of an interviewer perceived to be in a different socioeconomic and overall social situation than their own [51].

Conclusions

People experiencing homelessness can face many barriers when accessing digital technologies, including lack of appropriate devices, lack of operating infrastructure (eg, free Wi-Fi hotspots), some blind spots regarding digital skills, and a general lack of interest due to the prioritization of other basic life-supporting drives. However, in spite of all these barriers, our previous research identified a digitally engaged homeless subgroup in Budapest, Hungary, whose behaviors, usage, and access patterns were mapped in this study [30].

We found that the majority of participants (7/10) possessed a smartphone and used the often scarce pool of free Wi-Fi and the infrastructural capabilities of the shelters. Based on their articulated needs, various policy recommendations might be formulated for telephone companies and government agencies or support services. Telephone companies may consider subsidy programs to support mobile ownership and data services for this vulnerable population, as well as specific discount packages and more publicly available recharge options, as these would greatly support this group that is often in crisis and need. Government agencies may consider strengthening the infrastructural background of shelters and making free Wi-Fi accessibility an option in more public places, such as busses and piazza places, which could greatly reduce the access issues of this population. Institutional aid for accessing services and

XSL•FO
RenderX

digital tools may also offer a viable option for people experiencing homelessness. A higher digital accessibility of an institution in terms of both infrastructure and digital literacy is associated with a greater likelihood of an increase in the number of digitally engaged people experiencing homelessness.

In terms of usage patterns, digitally engaged people experiencing homelessness use digital tools as an alternative information point beyond mainstream health care channels, which gives them access to check information originating from mainstream health care personnel and to seek out complementary and alternative medical solutions. These might be related to low trust in mainstream health care solutions, which might be enhanced through appropriately tailored comprehensive digital health programs. These programs could include awareness raising programs on trusted online health information sources, digital literacy and health literacy enhancing programs, and other programs to enhance their general trust in evidence-based health and the health care system.

Our most important finding is that digitally engaged homeless individuals have an aptitude for technology, and they are ready and eager to share their knowledge with their peers. This could elevate them to the role of a mediator between their peers and any potential comprehensive digital health program. Digitally engaged individuals have the trust of their peers, recognize the benefits of digital technology, and are able to provide meaningful help in technology- and usage-related issues. Thus, with appropriate training, they might become tutors for upskilling people experiencing homelessness, building a bridge between their peers and digital technologies as well as digital health ecosystems. These well-informed technologically able peers might also help enhance trust in the general health care system if their peer-to-peer support could be steered toward peer-to-peer recommendations of trusted health information sources via a specific institutional program.

Overall, our previous research showed that digital health services have great promise in community shelters for managing and preventing health issues [29,30], and this study found that digitally engaged individuals might be important for the success of such services.

Acknowledgments

This research project was supported by the Hungarian National Research, Development and Innovation Office (OTKA-FK 134372).

Conflicts of Interest

None declared.

Multimedia Appendix 1
COREQ (Consolidated Criteria for Reporting Qualitative Research) checklist.
[PDF File (Adobe PDF File), 126 KB - humanfactors_v11i1e55415_app1.pdf]

Multimedia Appendix 2
Interview guide.
[PDF File (Adobe PDF File), 57 KB - humanfactors_v11i1e55415_app2.pdf]

References

1. The mobile economy 2023. GSMA. URL: https://www.gsma.com/solutions-and-impact/connectivity-for-good/mobile-economy/wp-content/uploads/2023/03/270223-The-Mobile-Economy-2023.pdf [accessed 2023-11-02]
2. Kemp S. Digital 2023 october global statshot report. DataReportal. URL: https://datareportal.com/reports/digital-2023-october-global-statshot [accessed 2024-04-30]
3. Sieck CJ, Sheon A, Ancker JS, Castek J, Callahan B, Siefer A. Digital inclusion as a social determinant of health. NPJ Digit Med 2021 Mar 17;4(1):52 [FREE Full text] [doi: 10.1038/s41746-021-00413-8] [Medline: 33731887]
4. Global strategy on digital health 2020-2025. World Health Organization. 2021. URL: https://www.who.int/publications-detail-redirect/9789240020924 [accessed 2024-04-30]
5. Ziebland S, Hyde E, Powell J. Power, paradox and pessimism: On the unintended consequences of digital health technologies in primary care. Soc Sci Med 2021 Nov;289:114419. [doi: 10.1016/j.socscimed.2021.114419] [Medline: 34619631]
6. van Kessel R, Hrzic R, O'Nuallain E, Weir E, Wong BLH, Anderson M, et al. Digital health paradox: International policy perspectives to address increased health inequalities for people living with disabilities. J Med Internet Res 2022 Feb 22;24(2):e33819 [FREE Full text] [doi: 10.2196/33819] [Medline: 35191848]
7. Sanders R. Digital inclusion, exclusion and participation. ESSS Outline. 2020. URL: https://www.iriss.org.uk/resources/esss-outlines/digital-inclusion-exclusion-and-participation [accessed 2024-04-30]
8. Menendez ME, Moverman MA, Puzzitiello RN, Pagani NR, Ring D. The telehealth paradox in the neediest patients. J Natl Med Assoc 2021 Jun;113(3):351-352 [FREE Full text] [doi: 10.1016/j.jnma.2020.09.144] [Medline: 33092857]

XSL•FO
RenderX

9. WHO Regional Office for South-East Asia. Social determinants of health. World Health Organization. 2008. URL: https://iris.who.int/handle/10665/206363 [accessed 2024-04-30]
10. Lathrop B. Moving toward health equity by addressing social determinants of health. Nurs Womens Health 2020 Feb;24(1):36-44. [doi: 10.1016/j.nwh.2019.11.003] [Medline: 31911097]
11. Braveman P, Gottlieb L. The social determinants of health: it's time to consider the causes of the causes. Public Health Rep 2014;129 Suppl 2(Suppl 2):19-31 [FREE Full text] [doi: 10.1177/00333549141291S206] [Medline: 24385661]
12. Steptoe A, Marmot M. The role of psychobiological pathways in socio-economic inequalities in cardiovascular disease risk. Eur Heart J 2002 Jan;23(1):13-25. [doi: 10.1053/euhj.2001.2611] [Medline: 11741358]
13. Cockerham WC, Hamby BW, Oates GR. The social determinants of chronic disease. Am J Prev Med 2017 Jan;52(1S1):S5-S12 [FREE Full text] [doi: 10.1016/j.amepre.2016.09.010] [Medline: 27989293]
14. Schreiter S, Bermpohl F, Krausz M, Leucht S, Rössler W, Schouler-Ocak M, et al. The prevalence of mental illness in homeless people in Germany. Dtsch Arztebl Int 2017 Oct 06;114(40):665-672 [FREE Full text] [doi: 10.3238/arztebl.2017.0665] [Medline: 29070426]
15. Zhang L, Norena M, Gadermann A, Hubley A, Russell L, Aubry T, et al. Concurrent disorders and health care utilization among homeless and vulnerably housed persons in Canada. J Dual Diagn 2018;14(1):21-31. [doi: 10.1080/15504263.2017.1392055] [Medline: 29494795]
16. Graffy P, McKinnon S, Lee G, Remington P. Life outside: A narrative ethnographic inquiry into the determinants of homelessness. J Poverty 2018 Dec 13;23(3):202-228. [doi: 10.1080/10875549.2018.1550133]
17. Bedmar MA, Bennasar-Veny M, Artigas-Lelong B, Salvà-Mut F, Pou J, Capitán-Moyano L, et al. Health and access to healthcare in homeless people: Protocol for a mixed-methods study. Medicine (Baltimore) 2022 Feb 18;101(7):e28816 [FREE Full text] [doi: 10.1097/MD.0000000000028816] [Medline: 35363172]
18. Aldridge RW, Story A, Hwang SW, Nordentoft M, Luchenski SA, Hartwell G, et al. Morbidity and mortality in homeless individuals, prisoners, sex workers, and individuals with substance use disorders in high-income countries: a systematic review and meta-analysis. Lancet 2018 Jan 20;391(10117):241-250 [FREE Full text] [doi: 10.1016/S0140-6736(17)31869-X] [Medline: 29137869]
19. van Dongen SI, van Straaten B, Wolf JRLM, Onwuteaka-Philipsen BD, van der Heide A, Rietjens JAC, et al. Self-reported health, healthcare service use and health-related needs: A comparison of older and younger homeless people. Health Soc Care Community 2019 Jul;27(4):e379-e388 [FREE Full text] [doi: 10.1111/hsc.12739] [Medline: 31020738]
20. Liu M, Hwang SW. Health care for homeless people. Nat Rev Dis Primers 2021 Jan 14;7(1):5. [doi: 10.1038/s41572-020-00241-2] [Medline: 33446661]
21. Odoh C, Vidrine JI, Businelle MS, Kendzor DE, Agrawal P, Reitzel LR. Health literacy and self-rated health among homeless adults. Health Behav Res 2019 Oct;2(4):13 [FREE Full text] [doi: 10.4148/2572-1836.1055] [Medline: 34164607]
22. Raven MC, Kaplan LM, Rosenberg M, Tieu L, Guzman D, Kushel M. Mobile phone, computer, and internet use among older homeless adults: Results from the HOPE HOME cohort study. JMIR Mhealth Uhealth 2018 Dec 10;6(12):e10049 [FREE Full text] [doi: 10.2196/10049] [Medline: 30530464]
23. VonHoltz LAH, Frasso R, Golinkoff JM, Lozano AJ, Hanlon A, Dowshen N. Internet and social media access among youth experiencing homelessness: Mixed-methods study. J Med Internet Res 2018 May 22;20(5):e184 [FREE Full text] [doi: 10.2196/jmir.9306] [Medline: 29789281]
24. Woods SS, Forsberg CW, Schwartz EC, Nazi KM, Hibbard JH, Houston TK, et al. The association of patient factors, digital access, and online behavior on sustained patient portal use: A prospective cohort of enrolled users. J Med Internet Res 2017 Oct 17;19(10):e345 [FREE Full text] [doi: 10.2196/jmir.7895] [Medline: 29042345]
25. Smith B, Magnani JW. New technologies, new disparities: The intersection of electronic health and digital health literacy. Int J Cardiol 2019 Oct 01;292:280-282 [FREE Full text] [doi: 10.1016/j.ijcard.2019.05.066] [Medline: 31171391]
26. Heaslip V, Richer S, Simkhada B, Dogan H, Green S. Use of technology to promote health and wellbeing of people who are homeless: A systematic review. Int J Environ Res Public Health 2021 Jun 25;18(13):6845 [FREE Full text] [doi: 10.3390/ijerph18136845] [Medline: 34202242]
27. Grech E, Raeburn T. Experiences of hospitalised homeless adults and their health care providers in OECD nations: A literature review. Collegian 2019 Feb;26(1):204-211 [FREE Full text] [doi: 10.1016/j.colegn.2018.04.004]
28. Győrffy Z, Békási S, Döbrössy B, Bognár V, Radó N, Morva E, et al. Exploratory attitude survey of homeless persons regarding telecare services in shelters providing mid- and long-term accommodation: The importance of trust. PLoS One 2022 Jan 6;17(1):e0261145 [FREE Full text] [doi: 10.1371/journal.pone.0261145] [Medline: 34990458]
29. Békási S, Girasek E, Győrffy Z. Telemedicine in community shelters: possibilities to improve chronic care among people experiencing homelessness in Hungary. Int J Equity Health 2022 Dec 17;21(1):181 [FREE Full text] [doi: 10.1186/s12939-022-01803-4] [Medline: 36528777]
30. Radó N, Girasek E, Békási S, Győrffy Z. Digital technology access and health-related internet use among people experiencing homelessness in Hungary: Quantitative survey. J Med Internet Res 2022 Oct 19;24(10):e38729 [FREE Full text] [doi: 10.2196/38729] [Medline: 36260379]

XSL•FO
RenderX

31. Győrffy Z, Radó N, Pölczman L, Sükösd A, Boros J. Creating work-life balance among physicians in the age of digitalization: the role of self-consciousness and communication - a qualitative study. BMC Health Serv Res 2023 Oct 24;23(1):1141 [FREE Full text] [doi: 10.1186/s12913-023-10101-w] [Medline: 37875908]
32. Malterud K, Siersma VD, Guassora AD. Sample size in qualitative interview studies: Guided by information power. Qual Health Res 2016 Nov;26(13):1753-1760. [doi: 10.1177/1049732315617444] [Medline: 26613970]
33. Steward JL, Stringfellow E, Schwartz S, Johnson-Roe NK, Holt CH, Pollio D, et al. Qualitative interviews with persons who have had experience with homelessness: Lessons Learned. 2010 Presented at: American Public Health Association 138th Annual Meeting And Exposition; November 6-10, 2010; Denver, CO.
34. Thomas DR. Feedback from research participants: are member checks useful in qualitative research? Qualitative Research in Psychology 2016 Aug 02;14(1):23-41. [doi: 10.1080/14780887.2016.1219435]
35. Braun V, Clarke V. Using thematic analysis in psychology. Qual Res Psychol 2006 Jan;3(2):77-101. [doi: 10.1191/1478088706qp063oa]
36. Bruce CR, Harrison P, Nisar T, Giammattei C, Tan NM, Bliven C, et al. Assessing the impact of patient-facing mobile health technology on patient outcomes: Retrospective observational cohort study. JMIR Mhealth Uhealth 2020 Jun 26;8(6):e19333 [FREE Full text] [doi: 10.2196/19333] [Medline: 32589161]
37. Ibrahim MS, Mohamed Yusoff H, Abu Bakar YI, Thwe Aung MM, Abas MI, Ramli RA. Digital health for quality healthcare: A systematic mapping of review studies. Digit Health 2022;8:20552076221085810 [FREE Full text] [doi: 10.1177/20552076221085810] [Medline: 35340904]
38. Uchino BN, Cacioppo JT, Kiecolt-Glaser JK. The relationship between social support and physiological processes: a review with emphasis on underlying mechanisms and implications for health. Psychol Bull 1996 May;119(3):488-531. [doi: 10.1037/0033-2909.119.3.488] [Medline: 8668748]
39. Bensken WP, Krieger NI, Berg KA, Einstadter D, Dalton JE, Perzynski AT. Health status and chronic disease burden of the homeless population: An analysis of two decades of multi-institutional electronic medical records. J Health Care Poor Underserved 2021;32(3):1619-1634 [FREE Full text] [doi: 10.1353/hpu.2021.0153] [Medline: 34421052]
40. Vallesi S, Tuson M, Davies A, Wood L. Multimorbidity among people experiencing homelessness-Insights from primary care data. Int J Environ Res Public Health 2021 Jun 16;18(12):6498 [FREE Full text] [doi: 10.3390/ijerph18126498] [Medline: 34208580]
41. Baggett TP, Liauw SS, Hwang SW. Cardiovascular disease and homelessness. J Am Coll Cardiol 2018 Jun 05;71(22):2585-2597 [FREE Full text] [doi: 10.1016/j.jacc.2018.02.077] [Medline: 29852981]
42. Braun E, Gazdag G. Pszichiátriai zavarok előfordulása hajléktalanok között [Prevalence of psychiatric disorders in homeless population]. Psychiatria Hungarica 2015;30(1):60-67 [FREE Full text]
43. Fazel S, Khosla V, Doll H, Geddes J. The prevalence of mental disorders among the homeless in western countries: systematic review and meta-regression analysis. PLoS Med 2008 Dec 02;5(12):e225 [FREE Full text] [doi: 10.1371/journal.pmed.0050225] [Medline: 19053169]
44. Tünde B. Mentális zavarral, hajléktalanságban élő kliensek ellátása. Budapest Methodological Centre of Social Policy and Its Institutions. 2022. URL: https://tinyurl.com/3myhs8je [accessed 2024-04-30]
45. McCarty D, Argeriou M, Huebner RB, Lubran B. Alcoholism, drug abuse, and the homeless. Am Psychol 1991 Nov;46(11):1139-1148. [doi: 10.1037//0003-066x.46.11.1139] [Medline: 1772151]
46. Győrffy Z, Döbrössy B, Radó N, Boros J, Békási S. State of vulnerable populations in the techquity framework in Hungary. Front Public Health 2023;11:1215325 [FREE Full text] [doi: 10.3389/fpubh.2023.1215325] [Medline: 37483940]
47. Harris J. The digitization of advice and welfare benefits services: re-imagining the homeless user. Housing Studies 2019 May 09;35(1):143-162. [doi: 10.1080/02673037.2019.1594709]
48. Miler JA, Carver H, Foster R, Parkes T. Provision of peer support at the intersection of homelessness and problem substance use services: a systematic 'state of the art' review. BMC Public Health 2020 May 07;20(1):641 [FREE Full text] [doi: 10.1186/s12889-020-8407-4] [Medline: 32381086]
49. Expanding peer support roles in homeless services delivery: A toolkit for service providers. Homeless and Housing Resource Center. 2023. URL: https://hhrctraining.org/system/files/paragraphs/download-file/file/2023-08/HHRC_Peer_Support_Toolkit-508.pdf [accessed 2024-04-30]
50. Barker SL, Maguire N. Experts by experience: Peer support and its use with the homeless. Community Ment Health J 2017 Jul;53(5):598-612 [FREE Full text] [doi: 10.1007/s10597-017-0102-2] [Medline: 28176210]
51. Stefkovics A, Sik E. What drives happiness? The interviewer's happiness. J Happiness Stud 2022;23:2745-2762. [doi: 10.1007/s10902-022-00527-0]

Abbreviations

HCSOM: Hungarian Charity Service of the Order of Malta
WHO: World Health Organization

XSL•FO
RenderX

Edited by C Jacob; submitted 12.12.23; peer-reviewed by N Karnik, S Markham, G Sebastian; comments to author 19.02.24; revised version received 02.03.24; accepted 10.03.24; published 14.05.24.

Please cite as:
Radó N, Békási S, Győrffy Z
Health Technology Access and Peer Support Among Digitally Engaged People Experiencing Homelessness: Qualitative Study
JMIR Hum Factors 2024;11:e55415
URL: https://humanfactors.jmir.org/2024/1/e55415
doi:10.2196/55415
PMID:38743937

©Nóra Radó, Sándor Békási, Zsuzsa Győrffy. Originally published in JMIR Human Factors (https://humanfactors.jmir.org), 14.05.2024. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on https://humanfactors.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Digital Lifestyle Interventions for Young People With Mental Illness: A Qualitative Study Among Mental Health Care Professionals

Chelsea Sawyer[1], PhD; Rebekah Carney[2], PhD; Lamiece Hassan[1], PhD; Sandra Bucci[1], DClinPsy; John Sainsbury[2], MSc; Karina Lovell[3], PhD; John Torous[4], MD; Joseph Firth[1], PhD

[1]Division of Psychology and Mental Health, University of Manchester, Manchester, United Kingdom
[2]Greater Manchester Mental Health NHS Foundation Trust, Manchester, United Kingdom
[3]Division of Nursing, Midwifery & Social Work, University of Manchester, Manchester, United Kingdom
[4]Beth Israel Deaconness Medical Centre, Harvard Medical School, Boston, MA, United States

Corresponding Author:
Joseph Firth, PhD
Division of Psychology and Mental Health
University of Manchester
3.005 Jean Mcfarlane Building
Manchester, M13 9PL
United Kingdom
Phone: 44 161 306 7811
Email: Joseph.firth@manchester.ac.uk

Abstract

Background: Given the physical health disparities associated with mental illness, targeted lifestyle interventions are required to reduce the risk of cardiometabolic disease. Integrating physical health early in mental health treatment among young people is essential for preventing physical comorbidities, reducing health disparities, managing medication side effects, and improving overall health outcomes. Digital technology is increasingly used to promote fitness, lifestyle, and physical health among the general population. However, using these interventions to promote physical health within mental health care requires a nuanced understanding of the factors that affect their adoption and implementation.

Objective: Using a qualitative design, we explored the attitudes of mental health care professionals (MHCPs) toward digital technologies for physical health with the goal of illuminating the opportunities, development, and implementation of the effective use of digital tools for promoting healthier lifestyles in mental health care.

Methods: Semistructured interviews were conducted with MHCPs (N=13) using reflexive thematic analysis to explore their experiences and perspectives on using digital health to promote physical health in youth mental health care settings.

Results: Three overarching themes from the qualitative analysis are reported: (1) motivation will affect implementation, (2) patients' readiness and capability, and (3) reallocation of staff roles and responsibilities. The subthemes within, and supporting quotes, are described.

Conclusions: The use of digital means presents many opportunities for improving the provision of physical health interventions in mental health care settings. However, given the limited experience of many MHCPs with these technologies, formal training and additional support may improve the likelihood of implementation. Factors such as patient symptomatology, safety, and access to technology, as well as the readiness, acceptability, and capability of both MHCPs and patients to engage with digital tools, must also be considered. In addition, the potential benefits of data integration must be carefully weighed against the associated risks.

(JMIR Hum Factors 2024;11:e53406) doi:10.2196/53406

KEYWORDS

digital health; behavior change; mental health care professionals; physical health; lifestyle intervention; qualitative; thematic analysis; service optimization; mobile phone

XSL•FO
RenderX

Introduction

Background

People with "severe mental illness" (SMI), such as schizophrenia, bipolar disorder, and associated psychotic or mood disorders, experience poorer physical health outcomes, which negatively affects their well-being across the life course and reduces life expectancy by up to approximately 15 years [1-3]. To reduce health disparities, it is crucial to adopt a preventative approach and intervene early. Adolescence or young adulthood presents a key opportunity as this is when most enduring mental health conditions are first diagnosed [4]. Young people with SMI and those at risk of SMI exhibit signs of poor cardiometabolic health and are more likely to engage in behaviors that are detrimental to their physical health; yet, much of this risk is modifiable [4,5].

Mental health care professionals (MHCPs) play a crucial role in supporting the mental and physical health needs of people with psychosis. However, MHCPs face significant barriers to delivering physical health interventions in practice [6]. This includes inadequate time and training in delivering evidence-based physical health interventions, difficulty reaching people in rural or remote areas, financial implications of delivering face-to-face interventions (particularly one-to-one), and limited National Health Service (NHS) resources for implementation [6].

Given these barriers and the increasing demand on the NHS, there is a growing focus on digital lifestyle interventions (DLIs), for example, using smartphones and websites to provide low-cost, scalable, and flexible interventions to promote healthier lifestyles [6,7]. The delivery of DLIs will require behavior changes among MHCPs. One model that can explain behavior change is the Capability, Opportunity, and Motivation–Behavior (COM-B) model [8]. According to this model, MHCP capability and opportunity to perform the behavior will influence their motivation to use DLIs and impact their delivery of DLIs in mental health care (MHC) settings. Capability refers to whether a person has the psychological (knowledge) or physical (skills) capability to perform the behavior. Alongside capability, an individual must have the opportunity to perform the behavior, and this refers to both physical (this includes the environment where the behavior will be performed and resources such as money and time) and social (the behavior of others) opportunity. Both reflective (reflective processes such as beliefs, goals, and values) and automatic (habitual and emotional responses) motivation also influence our behavior. The COM-B model can be used to inform future interventions [8].

Previous research suggests that MHCPs see the benefits of physical activity interventions [9]. However, MHCPs report barriers to implementation such as concerns about patient motivation and safety and logistical concerns on behalf of the patient, such as having equipment, clothes, and space. MHCPs have also reported personal barriers such as low confidence and capability to deliver interventions, lack of time and resources, and the belief that MHC should be a priority [10].

It is likely that MHCP attitudes toward and perceived barriers to using DLIs in MHC settings will vary from those for in-person interventions. According to actor-network theory, technology is not simply a tool or passive instrument that humans use to accomplish their goals [11]. Instead, technology can shape human behavior by creating new opportunities, alleviating constraints, and providing affordances that shape the way in which people think, communicate, and interact. Previous research has found that MHCPs believe that digital tools that support patient self-management would change their own roles and responsibilities [12]. Numerous studies have shown that, while MHCPs see the potential benefits, they are concerned about issues of liability, harm to patients, and lack of training regarding using DLIs [12-14]. As the MHCP role primarily focuses on treating mental health difficulties [15], it is important to explore and compare MHCP beliefs about using digital health for managing symptoms versus delivering lifestyle interventions.

Objectives

Therefore, this study aimed to explore MHCP perspectives, including barriers to and facilitators of using DLIs in MHC settings, with a particular focus on young people. These insights will provide key considerations for the implementation and use of DLIs in MHC settings.

Methods

Study Design

A mixed methods design was used, including a web-based survey and qualitative interviews, to examine the attitudes of MHCPs toward digital health in young people's MHC. An overview of the aligned findings of the combined survey and interview components of the project has been presented elsewhere [15]. While the previous mixed methods analysis provides a foundation for the research presented in this paper, this paper focuses on presenting the results of an in-depth qualitative examination of all the interview data, presenting the subjective experiences and perspectives of MHCPs regarding digital technology. Using a reflexive thematic analysis, we offered a more comprehensive understanding of subjective factors affecting the barriers and implementation of DLIs in clinical practice beyond the scope of the previous descriptive results to present useful recommendations for facilitating uptake of novel technology in health care settings.

The inclusion criteria were NHS MHCPs (1) working with young adults aged 16 to 35 years with mental illness, including specialist mental health services or in the context of broader primary care, and (2) working with young adults with mental illness for at least 6 months. The exclusion criteria were MHCPs working primarily with eating disorders due to differing treatment needs on nutrition and exercise [16-18]. The COREQ (Consolidated Criteria for Reporting Qualitative Research) checklist was used to ensure a comprehensive and explicit report of the interview process (Multimedia Appendix 1).

Participants

All 13 participants were MHCPs working within NHS services with young adults aged 16 to 35 years, including specialist

XSL•FO
RenderX

mental health services. Purposive sampling was used to recruit potential participants of a variety of occupational backgrounds and years of experience. Participants were recruited through emailed flyers. While interview participants were not reimbursed or rewarded for taking part in the interview, survey respondents were offered the choice to enter a prize draw to win a £50 (US $62.62) voucher. It was made clear to participants that completing the interview did not increase their chances of winning.

Data Collection

Overview

Semistructured interviews were conducted remotely using Microsoft Teams (Microsoft Corp) and audio recorded with participant consent. The interviews lasted 26 to 79 minutes and followed a topic guide (Multimedia Appendix 2) developed with input from our Patient and Public Involvement group and research team. In total, 2 researchers (CS and JF) conducted the interviews, which consisted of questions about participants' experience using digital health, the potential use of mobile health and DLIs in MHC, barriers to integration and use, and ways to boost engagement. Interview guides were flexible, using prompts and open questions to encourage participants to talk in depth about their experiences. All interviews were recorded and transcribed verbatim. Participants were assigned pseudonyms to maintain anonymity.

Data Analysis

Interviews were analyzed using a reflexive thematic approach [19-21]. Reflexive thematic analysis involves the researcher reflecting on how their experiences, personal assumptions, and background shape their analysis and interpretation of the data [20]. An inductively orientated experiential approach underpinned by critical realism was used [20]. This means that our themes were generated from the interviewees' direct experiences and observations while also recognizing that their understanding of reality is shaped by social and cultural factors. Critical realism was used as it allows the researcher to analyze participants' experiences while allowing the analysis to be informed by theory. The COM-B model of behavior change [21], which proposes that behavior is defined by our capability, opportunity, and motivation, was used as a theoretical underpinning and a prespecified area of interest. The model was used to identify potential barriers to the implementation of DLIs for young people with mental health conditions and any potential solutions to overcome these barriers.

Thematic analysis is a systematic approach whereby patterns and common themes are identified to describe a data set and understand a phenomenon [19,20]. The 6-phase guidelines by Clarke and Braun [20] were used to guide the analysis. These phases are recursive: (1) transcripts were read and reread so that the researcher (CS) could become familiar with the data, (2) systemic line-by-line coding was conducted to identify common features in the data, (3) codes were reviewed to determine themes, (4) themes were reviewed by 3 researchers (JF, CS, and LH) for homogeneity and heterogeneity to ensure that they were distinctive and coherent, (5) themes were defined and names were generated, and (6) findings were reported.

A primarily inductive approach was adopted with the interviews, but a deductive approach was taken when examining the barriers and facilitators informed by the COM-B model (based on previous research). Interview extracts related to barriers to and facilitators of implementing DLIs were mapped to the components of the COM-B model, whereas an inductive approach was taken for the remaining data. At the time of analysis, the researcher was not familiar with the current literature on digital health and in particular in the context of mental health, allowing them to analyze the data without preconceived themes or experiences.

To reduce the risk of bias, all researchers were involved in the analysis through regular meetings to discuss codes and themes. Themes and subthemes were generated and finalized using the NVivo software (version 12; QSR International) and MindView (version 7.0; Matchware). The team discussed the themes until consensus was reached within the team. During manuscript writing, subthemes with overlap were combined to avoid repetition. The final theme structure presented in the manuscript was reviewed and agreed upon by all coauthors.

Reflexivity

The two researchers who conducted the interviews (JF and CS) do not work in the NHS and made this clear to interviewees who were NHS employees. Both researchers have experience interviewing people on a variety of sensitive topics (self-harm, cancer [CS], and mental health [JF and CS]). All authors have an interest in digital health and promoting physical health in MHC settings. First author CS personally uses digital health apps. These views and experiences may have influenced our analysis; therefore, author CS kept a reflexive journal throughout the study. The author routinely reflected in the journal during data collection and analysis to reduce the possibility of their personal experiences and beliefs biasing their interpretations. A potential influencing factor in the interviews could be attributed to age. Some of the interviewees remarked that they did not have the same familiarity with apps as the interviewer (CS); this assumption could have influenced the views and opinions that participants expressed to this author.

Due to the COVID-19 pandemic, the interviews were conducted remotely using Microsoft Teams. Most participants joined the interviews from their own homes, providing a private setting that potentially fostered comfort and openness. However, working from home could have affected their work mindset and introduced distractions, such as pets, deliveries, or background noises. During an interview held in a private room at an interviewee's workplace, a team member interrupted the participant. While this interruption may have had an impact, it did not seem to alter their perceived barriers, and they continued discussing barriers, including those related to NHS staff, possibly influenced by their senior role. The remote format of Microsoft Teams interviews might have resulted in the interviewer missing out on subtle body language and facial responses, especially in the case of one participant who opted to keep their camera off. Nonetheless, Microsoft Teams provided flexibility, enabling participants to join at their preferred times.

XSL•FO
RenderX

Patient and Public Involvement

The overall study protocol had patient and carer involvement to ensure that all materials were appropriate and the content discussed about patients was appropriate and meaningful. The topic guide was developed using lived experience input. RC is a research fellow at a research unit embedded within clinical services at the Greater Manchester Mental Health NHS Foundation Trust (JUICE Youth Mental Health Research Unit). JUICE consists of academics as well as current practicing clinicians such as ward managers, lead psychiatrists, therapists, physiotherapists, dietitians, occupational therapists, experts by experience, and carers. A weekly consultation is held on the Child and Adolescent Mental Health Services inpatient units, where the topic guide was discussed.

Ethical Considerations

Ethics approval was granted by the University of Manchester Research Ethics Committee (2020-10603-17104) and the Health Research Authority (288734). Participants were briefed on the purpose of the study, and written informed consent was obtained.

Results

Participant Characteristics

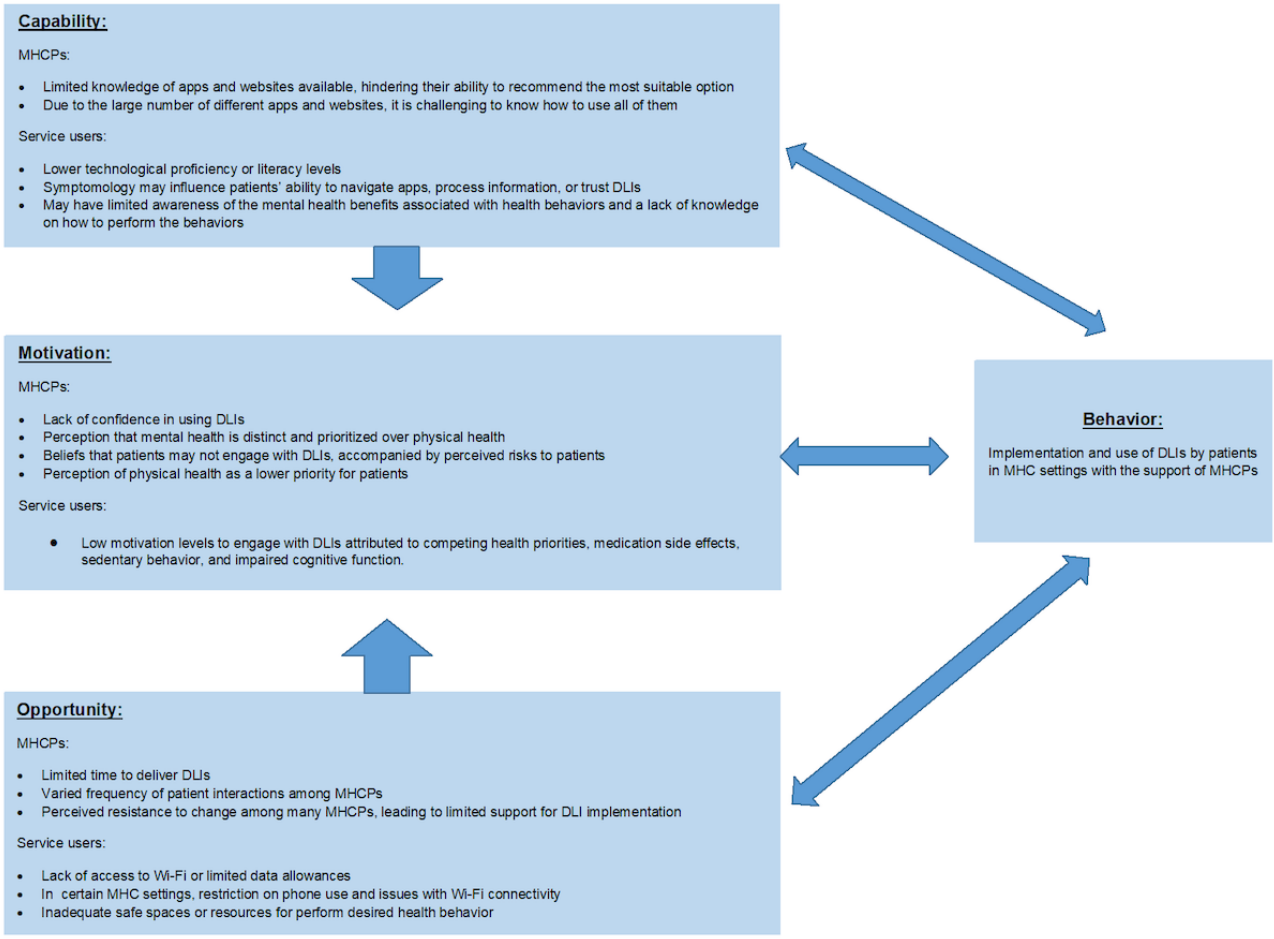
A total of 13 MHCPs were recruited from various MHC settings and roles, including MHC and research nurses (n=4, 31%), trainee psychiatrists (n=2, 15%), support workers (n=2, 15%), occupational therapists (n=1, 8%), physical health support workers (n=1, 8%), service managers (n=1, 8%), operational leads (n=1, 8%), and trainee advanced practitioners (n=1, 8%). Of the participants, 85% (11/13) were female, and 100% (13/13) were White British, with experience varying from at least 6 months to 20 years working in MHC service.

Analysis

Overview

Three main themes were identified: (1) motivation will affect implementation, (2) patients' readiness and capability, and (3) reallocation of staff roles and responsibilities. The themes and subthemes are described in Multimedia Appendix 3. Figure 1 presents how our findings also fit with the domains of the COM-B model [8].

Figure 1. Barriers to implementation of DLIs in MHC using the COM-B model_v2. COM-B: Capability, Opportunity, and Motivation–Behavior; DLI: digital lifestyle intervention; MHC: mental health care; MHCP: mental health care professional.



XSL•FO
RenderX

Theme 1: Motivation Will Affect Implementation

Overview

MHCP and patient motivation was perceived as the largest potential barrier to implementing DLIs in MHC settings. All interviewees felt that it would be important to emphasize the value and benefit of DLIs to patients, whereas acknowledging their potential risk is crucial for implementing DLIs in MHC settings:

> *I think some of them [MHCP], don't necessarily promote the apps, erm, because of the motivation...it can be quite problematic getting them [patients] to engage.* [INT5]

Subtheme 1: Individuals Are Motivated but Others Are Resistant

Most participants had a positive personal view of using digital health in MHC, in particular for young adults. However, some interviewees said that, although they were personally motivated to introduce and implement DLIs in the context of MHC delivery, their colleagues were reluctant to change, which posed a barrier to rolling out DLIs. MHCPs who are unaware of the benefits of DLIs or, for example, who are accustomed to established ways of working, were perceived by interviewees to be more resistant to adopting new practices:

> *...lots of people don't like change do they, and I think whenever you try and do anything new they'll be somebody that, will have something to say about it. I can't think of any specific kind of negative view of it, other than that, that just, traditional view that, oh it won't work, they won't use it.* [INT10]

Furthermore, several participants expressed concerns about time constraints and competing priorities faced by MHCPs. With an already demanding workload, interviewees said that MHCPs tend to prioritize mental health–focused care over physical health approaches or interventions. Interestingly, most people felt that their role was not best suited to the introduction of DLIs due to limited interactions with patients compared to their care coordinators despite being supportive of DLIs and their benefit more generally:

> *...it's [digital interventions] probably one of those things it's quite a good idea for the whole team to have an awareness of, but maybe, erm, you know, particularly care coordinators who are having the most contact with service users.* [INT9]

Subtheme 2: Patients Have Other Priorities

Interviewees perceived low motivation among young adults as a common barrier to engaging with DLIs. Low motivation was often attributed to medication side effects, sedentary behavior, and impaired cognitive function. Many interviewees perceived physical health to be a low priority for patients and that patients' main priority during their involvement with an MHC team was their mental health:

> *I think it depends on where they're at within their illness and how engaged they are with treatment and especially during those initial phases, it can be quite problematic, erm, and getting them to...engage.* [INT5]

Some interviewees felt that it was important to highlight the link between patients' physical and mental health and how changes to behavior could lead to changes in medication or treatment as well as improving physical and mental health:

> *Well talk about how, erm, maybe weight changes affect their mental health, how their medication has, the amount of medication that they've had has changed, how their, erm, you know, their diabetes diagnosis was reversed because they engaged in exercise and speak to them about what significance it would be to them, it's not about a size 8 jeans, it's about, I don't need to take as much Clozapine, and plus if I don't take as much Clozapine at night, then in a morning I'm not as knackered.* [INT5]

Interviewees also felt that the service setting might influence the uptake of DLIs. For example, young adults in inpatient settings may be hesitant to change due to the various restrictions in place, such as limiting takeaways or access to unhealthier food. DLIs that provided young adults with something such as promoting physical activity in inpatient settings were viewed as acceptable by interviewees. However, DLIs viewed as restrictive, such as those concerning diet or smoking, were assumed as not being well received by patients:

> *...inpatient mental health unit, erm, you've had so many bits of your identity taken away from you, in terms of like being able to access outside, and, I think, when you do try and have those conversations about, you know, well try and eat a bit healthier, it's gotten very angry really quickly because there's been so much of their liberties taken away, the fact that we try and take away the little things that they do enjoy like, staying in bed, erm, eating junk food, smoking, it, it, yeah, it can be quite a difficult subject.* [INT11]

Therefore, these preassumptions regarding what types of DLIs young adults are resistant to may reduce MHCP enthusiasm for undertaking the actions required to implement specific DLIs in inpatient settings. One MHC service that several MHCPs felt could work well to integrate DLIs in was early intervention services:

> *I think the young, younger people are more likely to use apps and You know when people first present to services, such as early intervention team. Uh, I think that would would work very, very well.* [INT1]

Subtheme 3: DLIs Need to Be Intuitive and Engaging

To overcome the perceived low motivation of patients to engage with DLIs, MHCPs felt that DLIs need to capture patients' attention and be engaging, which involved being visually appealing, interactive, and user-friendly, particularity due to the patients' age. Gamification, linking changes in behavior to patient-valued outcomes, and intuitive interfaces were perceived as being important in promoting engagement. Simple designs and usability were considered key, ensuring that both MHCPs and patients can navigate through the intervention:

XSL•FO
RenderX

> *...if you're focussing on young people, I think, I don't really see many barriers, if it's free, and, you know it's easy to use, I think it's just, it'd just be about that initial engagement, that initial kind of, them trying and it being good enough to keep them, er, interested.* [INT10]

> *...anything that's simple, straightforward and just easy to use would be probably the best starting point for now, for us [Mental Health are professionals].* [INT12]

> *I've never used gamified apps to be honest up, but yeah, it sounds great, It will make the younger people engage.* [INT2]

Several MHCPs shared the perspective that involving young adults with a wide spectrum of mental health conditions and literacy skills in the app design process is crucial to ensure both intuitiveness and engagement for patients:

> *...the important thing would be that if you were gonna design an app to, to genuinely have young people with a variety of mental health conditions, neurodevelopmental disorders, etc, all having input on what it looks like, how it works and how you engage with it.* [INT13]

Theme 2: Need to Consider Patients' Readiness and Capability to Use Digital Health

Subtheme 1: Patient Safety

A prominent theme was the need for patient benefit from digital health to outweigh potential risks or harm. Interviewees expressed the need to consider each patient and when was best to introduce DLIs. Initially, interviewees focused on what MHC settings would be the most suitable to introduce DLIs, with mixed views on the appropriateness of implementing DLIs in inpatient settings. On further reflection, the severity of the patients' symptoms rather than the MHC setting was considered most important when deciding on an appropriate time to introduce patients to DLIs:

> *...people with schizophrenia aren't unwell constantly, so you would use it and if they started to become paranoid or unwell, that's probably when you'd be able to just say, let's just remove it.* [INT12]

Concerns for patient safety were raised if DLIs required the young person to self-monitor physical activity or health data. For example, there were concerns that patients might misunderstand and assume that their MHC team was also monitoring their data, potentially leading them to not inform their MHC team of important changes. Alternatively, if health data were integrated and monitored by their MHC team, there were concerns that important data such as irregular heartbeat or reduced smoking while on clozapine could be missed, which could lead to harm:

> *I'm just thinking about difficulty I'd, I'd, feel terrible. If I had somebody on my caseload as a care coordinator and this information was there and I didn't pick up on it and I didn't notice that and then something happened, it's. Then there's a kind of risk factor there of Case negligence, maybe potentially, and who's going to be overseeing that.* [INT1]

Interviewees were also concerned that DLIs could worsen young adults' mental health symptoms, including paranoia, particularly for those who used phones or wearable devices. MHCPs articulated concerns that the tracking of patients' location and behavior via these wearable devices and phones could intensify patients' anxieties regarding perceived surveillance:

> *Consideration for people who could be psychotic, paranoid, suspicious, you know, and whether that might increase some of their Symptomatology, illusions or paranoia. You know if they were wearing a watch. For instance, knowing that I had access to that information.* [INT1]

MHCPs expressed concerns about potential risks that DLIs could pose regarding social interactions for young adults with psychosis. They were worried about inaccurate or harmful advice, exploitation, or negative interactions. Therefore, ensuring monitoring and moderation of social interactions within DLIs was necessary. However, interviewees also recognized the benefits of social support in boosting engagement:

> *...you would need some level of moderation in that community, 'cos you don't want unhelpful comments and views and, and unfortunately with online systems you get a lot of that. erm, because people are anonymous...there could be like a positive element of that and it could increase engagement.* [INT13]

Subtheme 2: Patients' Capability and Opportunity to Use Digital Technology

Digital health was perceived to be acceptable in MHC, especially among younger patients who were likely to have higher smartphone use. However, symptoms (eg, paranoia), reduced cognitive abilities (partly due to medication), and limited technological skills were seen as barriers to patients using digital technology:

> *A lot of the younger people are, erm [have technical skills], but, but also even, even the younger people when they're unwell, they have information processing problems.* [INT6]

Some MHCPs commented that paranoia may also hinder engagement. To address this, collecting minimal personal data through apps or websites was recommended:

> *I think just, the simpler the better maybe, you know, not having to gather as, as much information, just going off, you know, erm, individuals who, who can paranoid or, or the barriers, I think it, just something that's simple, that's easy, you know, you're not having to put a lot of data in.* [INT7]

Access was a potential barrier as some patients may lack access to phones, data, or apps, especially during periods of psychosis or as inpatients due to restrictions or poor Wi-Fi connections. Interviewees suggested that data-free apps that synchronize to Wi-Fi or apps with lower data requirements might overcome this barrier. To reduce the digital divide, interviewees said that services (ie, the NHS) should provide smartphones, pay for

XSL•FO
RenderX

subscription costs or data network charges, or provide wearables to ensure that recommended DLIs can be accessed fairly:

> *If something's gonna be effective and there's evidence base into it, I don't think that people should [pay], if it's about health and it's actually gonna reduce our costs in the long term, then it should be free.* [INT13]

Theme 3: Integrating Digital Health Will Require the Reallocation of Staff Roles and Responsibilities

Subtheme 1: Technology Changes Our Roles and Responsibilities

Despite attitudes being largely positive, there were differences in interviewees' perspectives on the impact that implementing DLIs into routine care would have on MHCP workload. Some interviewees who felt that their role should mainly focus on the mental health needs of a patient believed that DLIs would increase their workload. They suggested that the care coordinator (case manager) role would be best placed to implement DLIs as they have more patient contact and involvement and, therefore, have the time and a preexisting relationship:

> *...particularly care coordinators who are having the most contact with service users.* [INT9]

In contrast, interviewees in senior or physical health–focused professions viewed digital health as a way to reduce staff workload and enable frequent physical health monitoring (eg, remote blood pressure monitoring), resulting in better patient care; these interviewees believed that it was everyone's role and responsibility to implement DLIs in MHC:

> *...anything that's from that, the core components of the physical health check that they can input, would save everybody a lot of time and, erm, effort and money and it would also make it much more up to date.* [INT10]

Interviewees expressed concerns that the use of DLIs in MHC could affect their interactions with patients. Some interviewees believed that building a strong relationship with patients was crucial and DLIs entail a loss of face-to-face nuances, resulting in difficulties detecting physical symptoms, which in turn would be detrimental to patients' mental health. While some interviewees felt that DLIs may lead to people "becoming more isolated" (INT3), others felt that they would improve access to services for patients, particularly those with social anxieties and those who, due to their younger age, are more comfortable interacting digitally:

> *...with our cohort who might socially find things challenging and difficult actually a screen is quite familiar to them, so they will often prefer that.* [INT8]

Subtheme 2: MHCPs Will Need to Acquire Additional Skills

Interviewees said that staff involvement was crucial for successful implementation of DLIs in MHC settings. However, interviewees recognized that not all staff members have the necessary skills or knowledge to do this. Tailored training, including interactive sessions and MHCPs using the apps themselves, was recommended to enhance confidence, motivation, and psychological capability to use and recommend digital technology in clinical settings:

> *There's always room for training I find that interesting to go through it and if I was clinician, finding out what apps there are out there, how we can use them, how we can recommend.* [INT1]

One interviewee also felt that digital health education could be provided in health care degrees such as nursing and occupational therapy. Several MHCPs lacked awareness of available apps or websites. To address this, interviewees suggested that the NHS could provide a list of approved apps, improving trust and credibility while improving their knowledge and reducing guilt or personal blame in case of adverse events. Opinions differed on recommending apps without previous experience. Some expressed concerns about patient safety when recommending potentially ineffective or harmful apps, such as exercise apps that result in injury, whereas others felt comfortable if the apps were available to the general population:

> *...some people, if it's not NHS approved, might be a bit more nervous. I think that's a thing with the YouTube videos as well, like if it's just a person who's put together a, an exercise for somebody to follow, and it's not to do with the NHS, I think it does make people a bit more nervous about engaging people in that, but I personally have done, if, if I feel confident.* [INT11]

Subtheme 3: Who Is Responsible for Managing the Risk?

A concern regarding the implementation of DLIs was the issue of responsibility. All interviewees questioned who would be responsible if patients experienced negative outcomes as a result of using digital technology in the context of their health care. A particular concern was related to data monitoring, such as what data MHCPs should have access to and who should monitor them (patients or MHCPs). Interviewees expressed concerns about who was responsible for the oversight of digital data collection, especially if fluctuations in health or changes in behavior were missed. Interviewees wondered about their potential liability and described the guilt that they would feel if important changes in mental or physical health were missed. Some MHCPs strongly believed that data should never be integrated into NHS systems but patients could share and discuss their health with MHCPs if desired, empowering patients to take responsibility of their health:

> *I was thinking about the patient doing it for themselves rather than anybody having anybody sitting behind the scenes monitoring it.* [INT2]

On the other hand, some interviewees who routinely collected physical health data felt that, if DLIs collected data without MHCPs acknowledging or providing feedback in response, this could demotivate younger patients to use apps that track or record behavior:

> *I think for this kind of age group and even older, you need that, well done, you're doing well there, and not, relentless, I think having a barrier would be if there wasn't any kind of short-term goals that you could say, right we're making progress here, we're*

XSL•FO
RenderX

> *doing well, or this is what you need to work on.* [INT8]

> *...like the feedback, you know, 'cos a lot of the young people that, that we work with on the wards, they really, are seeking time with people and if you can provide that time that's quite focused on something and provide lots of positive reinforcement if something's gone really well, I think it might motivate people to, to get sustained use from an app like that, you probably need somebody who's on your side and really supporting you to use it well.* [INT11]

In addition, interviewees believed that young adults with mental health conditions may need support and advice from MHCPs in cases in which they implemented minor changes but did not observe any discernible outcomes, for example, if they made changes to their diet but did not lose weight. To overcome the burden of data monitoring and potential liability, one interviewee suggested implementing automated systems that notify clinicians of changes in heart rate, smoking behavior, and so on:

> *I prefer the idea of, I think it's a fantastic idea for them [young adult service users] to go away and when they come back and you say, your weight, you've been putting weight on, let's have a look on your app what you actually have been eating, so you can have those sort of discussions with them.* [INT6]

Discussion

Principal Findings

The aim of this study was to explore MHCP perspectives, including barriers to and facilitators of using DLIs to support young people with mental illness. To our knowledge, this is the first study to explore MHCP views on using and integrating digital health in MHC for young people. Overall, MHCPs felt that digital health care is acceptable when delivered alongside face-to-face care and has the potential to enhance the current care that patients receive. However, they also identified barriers to implementation, including staff and patient motivation and capability to deliver or use DLIs, concerns regarding patient safety, the digital divide, and the privacy of data.

Relevance to Previous Research and Existing Theory

Overview

Similar to previous research on lifestyle interventions [9,10,22], MHCPs expressed concerns with patient motivation and safety as well as time constraints and staff motivations and capability to deliver interventions. Notably, there were differences in concerns about resource availability as, while there were concerns about insufficient phone data interfering with DLIs, a lack of other resources (such as home exercise equipment, healthy eating ingredients, or staffing and clothing [9,22]) was not commonly raised as an issue in this study. The remote nature of DLIs raised concerns about missing important data, and key considerations for implementation included ensuring the credibility and trustworthiness of the apps or websites used in DLIs, prioritizing patient safety, and effective data monitoring.

In some cases, MHCPs may differentiate between the promotion of physical health and their core responsibilities. A decreased enthusiasm among MHCPs for addressing physical well-being may be tied to factors such as apprehension toward assuming personal responsibility and diminished patient motivation [10].

Our findings were in line with actor-network theory [11]—MHCPs did not perceive the implementation of DLIs as an additional resource to use. Instead, they felt that the implementation of DLIs would change their roles, bring new risks to patients, and affect rapport. However, a recent study exploring the views of patients with SMI on digital health found that they also believed that such technologies could change the relationship with MHCPs but in a positive way by empowering them to manage their health and providing a source of help other than their health care providers [18,23]. They also valued the ability to self-monitor and share their progress or behavior with their MHCPs to obtain additional support or positive feedback [23]. These results are also in line with those of broader digital health implementation studies suggesting a need for support for patients and clinicians as well as systems-related issues such as regulation, workflow, and safety [24]. Our findings also fit with the domains of the COM-B model [8] (Figure 1). Barriers identified from the interviews related to capability, opportunity, and motivation, along with potential solutions to overcome these barriers generated by the researchers, are presented in Table 1. The potential solutions were then mapped to potential intervention functions (broad categories to change behavior), and these potential barriers and solutions are discussed in more detail in the following sections.

XSL•FO
RenderX

Table 1. Potential problems and solutions for the implementation of digital lifestyle interventions (DLIs) in mental health care settings.

Domain and problem	Solution	Intervention function
Capability		
Patients may have poorer digital literacy and skills to use technology or phones (perceived as less of an issue for younger populations)	Instructions on how to effectively use phones and apps	Education
Symptoms that may limit patients' cognitive ability to use digital technology (eg, side effects of medication)	Introduce DLIs only when patients have the mental capacity to consent	Restriction
MHCPs[a] do not know which apps are available and effective and how to use them	Training on how to find and use apps and knowledge sharing within clinical teams	Training or education
Opportunity		
Patients not having access to phones, internet, or data	Provision of phones, payment for data and wearables, and use of apps that are usable offline and resynchronize or update once connected to Wi-Fi or data	Enablement or environmental restructuring
Cost of wearables, data, and app subscriptions	Provision of wearables, NHS[b] covering the cost of app subscriptions, and having an iPad or device that can be used by an entire inpatient ward or service	Enablement or environmental restructuring
Restrictions (eg, no phones and no space) and restricted use or functionality when in inpatient units	Having a shared device on the ward or having supervised access to use the app and being shown how to use it before discharge	Service provision
MHCPs do not have the time to deliver interventions	Peer coaches or digital navigators to help patients install apps and deal with issues	Environmental restructuring
Integrating health data to ensure patient safety	Automated notifications if there are high risks or behavior changes that need to be addressed (ie, changes in smoking on clozapine)	Enablement
Secure storage of data	Improvement in the current technology infrastructure of the NHS to allow app data to be securely stored on the system	Enablement
Motivation		
Patients having low motivation	To boost motivation, use rewards, games, self-monitoring of behaviors or health outcomes, and provision of feedback	Incentives
MHCPs lack confidence with digital technology	Training and education and having regular drop-in clinics to problem solve any issues	Training
MHCPs' competing priorities (focus on treating mental health)	When implementing DLIs, services will need to consider which MHCP roles would be best to implement this and could use individual or service-based targets regarding delivering DLIs	Incentives or coercion
MHCP beliefs that patients do not want to change behavior	Involve patients in the integration of DLIs; this will demonstrate that patients are willing to use DLIs	Persuasion
MHCPs do not see the benefit of digital interventions	Educating MHCPs on the benefit of treating physical and mental health together and how digital health can play a role in this	Education

[a]MHCP: mental health care professional.

[b]NHS: National Health Service.

Capability

MHCPs voiced uncertainties about their own confidence and ability to effectively deliver DLIs. According to our previous work, MHCPs have limited opportunities to use DLIs in their current role [15], and this is consistent with current literature [13,25]. This may have led MHCPs to perceive a lack of psychological capability (knowledge) to recommend and use apps. Therefore, appropriate training is required, such as interactive training sessions that allow MHCPs to trial various apps and websites, thereby boosting their self-efficacy (the belief in their ability to perform the behavior) and motivation [26]. Alternatively, some MHCPs suggested that participating in DLIs themselves could enhance their motivation. This notion is supported by recent research demonstrating that MHCPs who actively participated in exercise were more inclined to encourage its adoption among inpatients [10].

A lack of knowledge of apps available was perceived as a barrier. MHCPs felt that receiving a list of approved apps to recommend would reinforce the credibility and trustworthiness of the apps, remove personal liability, and improve their

XSL•FO
RenderX

knowledge about what is available. Interestingly, very few MHCPs were aware of organizations such as the Organisation for the Review of Care and Health Apps or SilverCloud, which provide a list of NHS-approved apps. Awareness of approved and endorsed apps could reduce the burden on MHCPs in selecting appropriate apps for patients. For example, the MINDapps database allows users to search for mental health apps using specific criteria and provides a description, a rating, and reviews of each app.

MHCPs emphasized the need for considering the clinical presentation of young adults with a mental health condition before the initiation of DLIs and continuously monitoring it should any clinical deterioration occur. They stressed the importance of evaluating symptom severity and diagnosis, with particular care taken for individuals with eating disorders (when recommending diet or physical activity interventions). The focus was on prioritizing patient safety and avoiding potential harm while acknowledging that symptom severity influences patient engagement and motivation regarding physical health interventions. In addition, patients may have limited technology skills or literacy and may need additional support or training on how to use apps [27]. Promisingly, previous research [28,29] has shown that individuals who are less familiar with DLIs can use them after minimal training or with support from peer coaches. Incorporating peer coaches to train and support patients could help lighten the workload burden on MHCPs and contribute to reducing the digital divide [30].

Opportunity

Competing interests and limited time, limited patient contact, and inadequate technology infrastructure in the NHS were perceived as barriers to implementing DLIs in MHC settings for MHCPs. This was in line with a recent study that found that MHC settings lacked effective integration of digital technologies [31]. The study also found significant differences among MHCPs regarding whether it falls under their role to implement digital MHC [31]. Therefore, ensuring successful integration of DLIs in MHC requires a thoughtful evaluation of their alignment with existing services, restrictions in inpatient settings, the optimal MHCP role or roles to deliver DLIs, and ways to minimize unnecessary burden on MHCPs.

One solution alongside improvements in technology infrastructure is new MHCP roles dedicated to implementing DLIs and monitoring patient data in MHC settings via the digital navigator pathway [30]. In addition to alleviating burdens, digital navigators could provide feedback to patients on their data. Berry et al [32] discovered that patients expressed a keen interest in engaging with MHCPs to review digitally collected outcome data related to their mental health. This notion was mirrored in this study, with MHCPs emphasizing the need for feedback on health data to gain deeper insights into causal links between behavior and health or to underscore the effectiveness of subtle changes that might take time to be realized more broadly. Although concerns were expressed about collecting and monitoring health data, there are several benefits to using near–real-time data—they can improve the quality of care received, detect health deterioration earlier, and reduce staff demands through more efficient monitoring [33-35].

In contrast with other research exploring MHCPs' views on digital health for the self-management of severe mental health conditions [32], we found that access to phones was not perceived as a barrier, likely due to the young age of the patient population (18-35 years) and increase in mobile phone ownership in the last decade [36]. However, limited data and app subscriptions were perceived as significant barriers. Staff felt that the NHS should provide phones and wearables and cover the cost of app subscriptions and data allowance. If the NHS is to cover app subscriptions and data allowance, the cost-effectiveness of DLIs needs to be considered to reduce the digital divide [30,37,38].

Motivation

In line with previous research, staff were reluctant to promote physical health [10,39,40] because it was perceived as not part of their role. MHCPs also lacked confidence in delivering DLIs and perceived that young people with mental illness have other priorities than their physical health and that risk of harm to patients may outweigh potential benefit. This perceived distinction between physical and mental health needs to be addressed through wider changes in education or service provision to highlight the need to take a holistic approach and treat mental and physical health concurrently [10]. Staff were particularly reluctant to implement DLIs for young people in inpatient settings, where the environment may be more unpredictable and access to technologies may be restricted in some cases. Furthermore, MHCPs felt that young people may be difficult to engage in this environment due to feeling out of control and to restrictions on movement [41]. However, introducing DLIs may provide patients with the autonomy to look after their own physical health when they feel out of control [42] as well as the potential therapeutic effects from physical activity [43]. To address MHCPs' concerns regarding professional liability and patient safety, clear guidelines and frameworks will need to be in place. MHCPs also stressed the significance of engaging both MHCPs and patients in the development and implementation of DLIs in MHC, believing that this collaborative approach can enhance overall buy-in from MHCPs and patients alike.

Strengths and Limitations

A strength of this research is the potential real-world impact on patient care by providing a rich understanding of MHCP experiences and clinical recommendations to implement digital health in MHC settings. Our sample comprised primarily White British and female individuals; this means that the findings may not be transferable across cultures, gender, and ethnicities. As with all qualitative research, this study was shaped by the researchers' personal experiences and views stemming from their own use of digital health and views on DLIs. A reflexive approach among the research team reflecting on the researchers' personal experiences, assumptions, and perspectives was used throughout the research process.

Recommendations for Implementation

We make 5 recommendations based on our findings. First, clear guidelines for recommending apps, handling data, and monitoring safety are needed. These guidelines should be

XSL•FO
RenderX

developed with stakeholders, patients, and MHCPs to ensure that their concerns and needs are addressed. Second, DLIs should be coproduced, to some extent, with their intended end users to ensure that they are appropriate, engaging, and user-friendly and with those who will be delivering the DLIs to ensure that MHCPs have the skills and these steps may lead to better engagement. While coproduction of DLIs could involve full cocreation of bespoke technologies, this may not always be required, and instead, coproducing the way in which an existing technology is implemented in MHC and provided to patients could be sufficient. Third, to boost MHCPs' psychological capability and skills in using digital technology, training should be provided. This training could be provided by colleagues who were involved in implementing DLIs to gain clinician buy-in or as part of their professional training, which could lead to increased adoption, better implementation, and improved outcomes for patients. Alternatively, the delivery of this training by patients could help dispel certain preconceived notions, such as the belief that service users may be unable or unwilling to effectively use apps. Fourth, MHCPs need to be made aware of evidence-based digital resources available to them. Finally, a patient-centered approach should be used when implementing DLIs that considers patients' current symptomatology, particularly regarding paranoia, and clear criteria should be set to reduce any potential risk.

Future Research

Areas for future consideration toward real-world implementation are (1) the provision of apps, data, or phones to reduce the digital divide; (2) the practical, ethical, and security issues regarding the collection and monitoring of health data; and (3) which MHCP role would be best to implement DLIs or whether a new role is required. First, bridging the digital divide is crucial, and we need to ensure equal opportunities for those with mental health conditions to engage with DLIs. Therefore, future work needs to determine the preventative cost and impact of providing DLIs in MHC settings, focusing on patients' safety, efficacy, the quality of care received, and the impact on patients' physical health and their use of other health services. Second, it is important to determine the parameters for data collection. This includes working to determine the optimal type and frequency of health or behavior data and how this can be applicable beyond research purposes and actually become clinically useful for improving patient outcomes. In addition, there is a need to address the ethical considerations surrounding data access in clinical settings, ensuring that patients are able to provide informed consent and outlining what happens when patients lose the capacity to provide informed consent. Finally, further efforts are required to establish whether MHCPs can adequately deliver DLIs in their current roles or whether the development of a new role is necessary to maximize the effectiveness of DLIs in MHC settings. In summary, future research is required to examine the sustainability and cost-effectiveness of the implementation of DLIs, the optimal times and means for introducing DLIs to patients, and how this can be tailored to suit the individual needs of people diagnosed with a mental health condition.

Conclusions

Implementing lifestyle interventions for individuals with SMI is imperative, and the incorporation of DLIs may overcome some of the barriers faced by in-person interventions. Alongside this, implementation during the early intervention period could present a pivotal opportunity for timely approaches to preventing physical comorbidities from arising. The findings from this study suggest that, while digital health has the potential to enhance MHC and the quality of care that patients receive, there are important concerns that MHCPs hold that need to be addressed when considering implementation. Efforts are required to work with patients, MHCPs, and other stakeholders to identify appropriate content and delivery of DLIs, along with the types and content of training required to facilitate their implementation in routine clinical practice.

Acknowledgments

The authors would like to thank all the mental health care professionals who participated and helped recruit participants for this study. This study was funded by the University of Manchester Presidential Fellowship (P123958) and a UK Research and Innovation Future Leaders Fellowship (MR/T021780/1). SB is funded by a research professorship (NIHR300794).

Data Availability

Data sharing is not applicable to this paper as no data sets were generated or analyzed during this study.

Authors' Contributions

JF, RC, JS, and SB developed the study design. JF supervised the study. CS and JF collated and analyzed the data. CS, LH, and JF discussed the themes. CS and JF wrote the manuscript. JF, SB, LH, RC, and JS provided substantial input throughout the development and writing of the manuscript.

Conflicts of Interest

JF has received honoraria and consultancy fees from Atheneum, Informa, Gillian Kenny Associates, Big Health, Nutritional Medicine Institute, ParachuteBH, Richmond Foundation, and Nirakara independent of this work. SB is the director and a shareholder of CareLoop Health Ltd, a spinout from the University of Manchester to develop and market digital solutions for remote monitoring using smartphones for mental health conditions (currently, schizophrenia and postnatal depression). JT is the editor of JMIR

XSL•FO
RenderX

Mental Health and is on the scientific board of Precision Mental Wellness; however, this is unrelated to this work. The remaining authors (CS, RC, and LH) declare no other conflicts of interest.

Multimedia Appendix 1
COREQ (Consolidated Criteria for Reporting Qualitative Research) checklist.
[DOCX File , 17 KB - humanfactors_v11i1e53406_app1.docx]

Multimedia Appendix 2
Interview schedule.
[DOCX File , 15 KB - humanfactors_v11i1e53406_app2.docx]

Multimedia Appendix 3
Theme summary table.
[DOCX File , 25 KB - humanfactors_v11i1e53406_app3.docx]

References

1. Firth J, Siddiqi N, Koyanagi A, Siskind D, Rosenbaum S, Galletly C, et al. The Lancet Psychiatry Commission: a blueprint for protecting physical health in people with mental illness. Lancet Psychiatry 2019 Aug;6(8):675-712. [doi: 10.1016/S2215-0366(19)30132-4] [Medline: 31324560]
2. Solmi M, Fornaro M, Ostinelli EG, Zangani C, Croatto G, Monaco F, et al. Safety of 80 antidepressants, antipsychotics, anti-attention-deficit/hyperactivity medications and mood stabilizers in children and adolescents with psychiatric disorders: a large scale systematic meta-review of 78 adverse effects. World Psychiatry 2020 Jun 11;19(2):214-232 [FREE Full text] [doi: 10.1002/wps.20765] [Medline: 32394557]
3. Perry BI, Stochl J, Upthegrove R, Zammit S, Wareham N, Langenberg C, et al. Longitudinal trends in childhood insulin levels and body mass index and associations with risks of psychosis and depression in young adults. JAMA Psychiatry 2021 Apr 01;78(4):416-425 [FREE Full text] [doi: 10.1001/jamapsychiatry.2020.4180] [Medline: 33439216]
4. Carney R, Cotter J, Bradshaw T, Firth J, Yung AR. Cardiometabolic risk factors in young people at ultra-high risk for psychosis: a systematic review and meta-analysis. Schizophr Res 2016 Feb;170(2-3):290-300 [FREE Full text] [doi: 10.1016/j.schres.2016.01.010] [Medline: 26794596]
5. Carney R, Firth J, Pedley R, Law H, Parker S, Lovell K. The clinical and behavioral cardiometabolic risk of children and young people on mental health inpatient units: a systematic review and meta-analysis. Gen Hosp Psychiatry 2021;70:80-97 [FREE Full text] [doi: 10.1016/j.genhosppsych.2021.03.007] [Medline: 33773375]
6. Weightman M. Digital psychotherapy as an effective and timely treatment option for depression and anxiety disorders: implications for rural and remote practice. J Int Med Res 2020 Jun;48(6):300060520928686 [FREE Full text] [doi: 10.1177/0300060520928686] [Medline: 32527170]
7. Stefanopoulou E, Lewis D, Taylor M, Broscombe J, Larkin J. Digitally delivered psychological interventions for anxiety disorders: a comprehensive review. Psychiatr Q 2019 Mar;90(1):197-215. [doi: 10.1007/s11126-018-9620-5] [Medline: 30488330]
8. Michie S, van Stralen MM, West R. The behaviour change wheel: a new method for characterising and designing behaviour change interventions. Implement Sci 2011 Apr 23;6:42 [FREE Full text] [doi: 10.1186/1748-5908-6-42] [Medline: 21513547]
9. Martland R, Gaughran F, Stubbs B, Onwumere J. Perspectives on implementing HIIT interventions for service users in inpatient mental health settings: a qualitative study investigating patient, carer and staff attitudes. J Affect Disord 2021 Mar 15;283:198-206. [doi: 10.1016/j.jad.2021.01.065] [Medline: 33561800]
10. Ball H, Yung A, Bucci S. Staff perspectives on the barriers and facilitators to exercise implementation in inpatient mental health services: a qualitative study. Ment Health Phys Act 2022 Mar;22:100452. [doi: 10.1016/j.mhpa.2022.100452]
11. Latour B. Reassembling the Social: An Introduction to Actor-Network-Theory. Oxford, UK: Oxford University Press; 2007.
12. Berry N, Bucci S, Lobban F. Use of the internet and mobile phones for self-management of severe mental health problems: qualitative study of staff views. JMIR Ment Health 2017 Nov 01;4(4):e52 [FREE Full text] [doi: 10.2196/mental.8311] [Medline: 29092809]
13. Sterling WA, Sobolev M, Van Meter A, Guinart D, Birnbaum ML, Rubio JM, et al. Digital technology in psychiatry: survey study of clinicians. JMIR Form Res 2022 Nov 10;6(11):e33676 [FREE Full text] [doi: 10.2196/33676] [Medline: 36355414]
14. Schmidt S, D'Alfonso S. Clinician perspectives on how digital phenotyping can inform client treatment. Acta Psychol (Amst) 2023 May;235:103886 [FREE Full text] [doi: 10.1016/j.actpsy.2023.103886] [Medline: 36921359]
15. Sawyer C, Hassan L, Sainsbury J, Carney R, Bucci S, Burgess H, et al. Using digital technology to promote physical health in mental healthcare: a sequential mixed-methods study of clinicians' views. Early Interv Psychiatry 2024 Feb;18(2):140-152. [doi: 10.1111/eip.13441] [Medline: 37318221]

XSL•FO
RenderX

16. Eikey EV, Reddy MC, Booth KM, Kvasny L, Blair JL, Li V, et al. Desire to be underweight: exploratory study on a weight loss app community and user perceptions of the impact on disordered eating behaviors. JMIR Mhealth Uhealth 2017 Oct 12;5(10):e150 [FREE Full text] [doi: 10.2196/mhealth.6683] [Medline: 29025694]
17. Eikey EV, Reddy MC. "It's definitely been a journey": a qualitative study on how women with eating disorders use weight loss apps. In: Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems. 2017 Presented at: CHI '17; May 6-11, 2017; Denver, CO p. 642-654 URL: https://dl.acm.org/doi/10.1145/3025453.3025591 [doi: 10.1145/3025453.3025591]
18. Torous J, Bucci S, Bell IH, Kessing LV, Faurholt-Jepsen M, Whelan P, et al. The growing field of digital psychiatry: current evidence and the future of apps, social media, chatbots, and virtual reality. World Psychiatry 2021 Oct 09;20(3):318-335 [FREE Full text] [doi: 10.1002/wps.20883] [Medline: 34505369]
19. Braun V, Clarke V. Using thematic analysis in psychology. Qual Res Psychol 2006 Jan;3(2):77-101. [doi: 10.1191/1478088706qp063oa]
20. Braun V, Clarke V. Thematic Analysis: A Practical Guide. Thousand Oaks, CA: SAGE Publications; 2021.
21. Braun V, Clarke V. Can I use TA? Should I use TA? Should I not use TA? Comparing reflexive thematic analysis and other pattern-based qualitative analytic approaches. Couns Psychother Res 2020 Oct 18;21(1):37-47. [doi: 10.1002/capr.12360]
22. Happell B, Platania-Phung C, Scott D. Placing physical activity in mental health care: a leadership role for mental health nurses. Int J Ment Health Nurse 2011 Oct;20(5):310-318. [doi: 10.1111/j.1447-0349.2010.00732.x] [Medline: 21896121]
23. Williams J, Gillis C, Williams J, Gillett G, McGrath R, Ang K, et al. Exploring the use of digital interventions by people with severe mental illness to support their physical health: a mixed methods study. Issues Ment Health Nurs 2024 Jan;45(1):9-26. [doi: 10.1080/01612840.2023.2279207] [Medline: 38190422]
24. van Kessel R, Roman-Urrestarazu A, Anderson M, Kyriopoulos I, Field S, Monti G, et al. Mapping factors that affect the uptake of digital therapeutics within health systems: scoping review. J Med Internet Res 2023 Jul 25;25:e48000 [FREE Full text] [doi: 10.2196/48000] [Medline: 37490322]
25. Mahmoud K, Jaramillo C, Barteit S. Telemedicine in low- and middle-income countries during the COVID-19 pandemic: a scoping review. Front Public Health 2022 Jun 22;10:914423 [FREE Full text] [doi: 10.3389/fpubh.2022.914423] [Medline: 35812479]
26. Bandura A. Self-efficacy: toward a unifying theory of behavioral change. Psychol Rev 1977 Mar;84(2):191-215. [doi: 10.1037//0033-295x.84.2.191] [Medline: 847061]
27. Spanakis P, Wadman R, Walker L, Heron P, Mathers A, Baker J, et al. Measuring the digital divide among people with severe mental ill health using the essential digital skills framework. Perspect Public Health 2024 Jan 05;144(1):21-30 [FREE Full text] [doi: 10.1177/17579139221106399] [Medline: 35929589]
28. Olmos-Ochoa TT, Niv N, Hellemann G, Cohen AN, Oberman R, Goldberg R, et al. Barriers to participation in web-based and in-person weight management interventions for serious mental illness. Psychiatr Rehabil J 2019 Sep;42(3):220-228 [FREE Full text] [doi: 10.1037/prj0000363] [Medline: 31081651]
29. Young AS, Cohen AN, Goldberg R, Hellemann G, Kreyenbuhl J, Niv N, et al. Improving weight in people with serious mental illness: the effectiveness of computerized services with peer coaches. J Gen Intern Med 2017 Apr 7;32(Suppl 1):48-55 [FREE Full text] [doi: 10.1007/s11606-016-3963-0] [Medline: 28271427]
30. Wisniewski H, Gorrindo T, Rauseo-Ricupero N, Hilty D, Torous J. The role of digital navigators in promoting clinical care and technology integration into practice. Digit Biomark 2020 Nov 26;4(Suppl 1):119-135 [FREE Full text] [doi: 10.1159/000510144] [Medline: 33442585]
31. LaMonica HM, Iorfino F, Lee GY, Piper S, Occhipinti JA, Davenport TA, et al. Informing the future of integrated digital and clinical mental health care: synthesis of the outcomes from project synergy. JMIR Ment Health 2022 Mar 09;9(3):e33060 [FREE Full text] [doi: 10.2196/33060] [Medline: 34974414]
32. Berry N, Lobban F, Emsley R, Bucci S. Acceptability of interventions delivered online and through mobile phones for people who experience severe mental health problems: a systematic review. J Med Internet Res 2016 May 31;18(5):e121 [FREE Full text] [doi: 10.2196/jmir.5250] [Medline: 27245693]
33. Torous J, Bucci S, Bell IH, Kessing LV, Faurholt-Jepsen M, Whelan P, et al. The growing field of digital psychiatry: current evidence and the future of apps, social media, chatbots, and virtual reality. World Psychiatry 2021 Oct 09;20(3):318-335 [FREE Full text] [doi: 10.1002/wps.20883] [Medline: 34505369]
34. Marzano L, Bardill A, Fields B, Herd K, Veale D, Grey N, et al. The application of mHealth to mental health: opportunities and challenges. Lancet Psychiatry 2015 Oct;2(10):942-948. [doi: 10.1016/s2215-0366(15)00268-0]
35. Macrynikola N, Nguyen N, Lane E, Yen S, Torous J. The digital clinic: an innovative mental health care delivery model utilizing hybrid synchronous and asynchronous treatment. NEJM Catal Innov Care Deliv 2023 Aug 16;4(9). [doi: 10.1056/cat.23.0100]
36. Firth J, Cotter J, Torous J, Bucci S, Firth JA, Yung AR. Mobile phone ownership and endorsement of "mHealth" among people with psychosis: a meta-analysis of cross-sectional studies. Schizophr Bull 2016 Mar 22;42(2):448-455 [FREE Full text] [doi: 10.1093/schbul/sbv132] [Medline: 26400871]

XSL•FO
RenderX

37. Sylvia LG, Faulkner M, Rakhilin M, Amado S, Gold AK, Albury EA, et al. An online intervention for increasing physical activity in individuals with mood disorders at risk for cardiovascular disease: design considerations. J Affect Disord 2021 Aug 01;291:102-109. [doi: 10.1016/j.jad.2021.04.094] [Medline: 34029880]
38. Wisniewski H, Torous J. Digital navigators to implement smartphone and digital tools in care. Acta Psychiatr Scand 2020 Apr 26;141(4):350-355 [FREE Full text] [doi: 10.1111/acps.13149] [Medline: 31930477]
39. Kohn L, Christiaens W, Detraux J, De Lepeleire J, De Hert M, Gillain B, et al. Barriers to somatic health care for persons with severe mental illness in Belgium: a qualitative study of patients' and healthcare professionals' perspectives. Front Psychiatry 2021 Jan 26;12:798530 [FREE Full text] [doi: 10.3389/fpsyt.2021.798530] [Medline: 35153863]
40. Happell B, Scott D, Platania-Phung C. Perceptions of barriers to physical health care for people with serious mental illness: a review of the international literature. Issues Ment Health Nurs 2012 Nov 12;33(11):752-761. [doi: 10.3109/01612840.2012.708099] [Medline: 23146009]
41. Tully SM, Bucci S, Berry K. "My life isn't my life, it's the systems": a qualitative exploration of women's experiences of day-to-day restrictive practices as inpatients. J Psychiatr Ment Health Nurs 2023 Feb 18;30(1):110-122 [FREE Full text] [doi: 10.1111/jpm.12855] [Medline: 35771190]
42. Szinay D, Perski O, Jones A, Chadborn T, Brown J, Naughton F. Perceptions of factors influencing engagement with health and well-being apps in the United Kingdom: qualitative interview study. JMIR Mhealth Uhealth 2021 Dec 16;9(12):e29098. [doi: 10.2196/29098] [Medline: 34927597]
43. Bull ER, McCleary N, Li X, Dombrowski SU, Dusseldorp E, Johnston M. Interventions to promote healthy eating, physical activity and smoking in low-income groups: a systematic review with meta-analysis of behavior change techniques and delivery/context. Int J Behav Med 2018 Dec 12;25(6):605-616 [FREE Full text] [doi: 10.1007/s12529-018-9734-z] [Medline: 30003476]

Abbreviations

COM-B: Capability, Opportunity, and Motivation–Behavior
COREQ: Consolidated Criteria for Reporting Qualitative Research
DLI: digital lifestyle intervention
MHC: mental health care
MHCP: mental health care professional
NHS: National Health Service
SMI: severe mental illness

Edited by A Kushniruk, A Parush; submitted 17.11.23; peer-reviewed by B Lambe, N Josifovski, C Baxter; comments to author 11.01.24; revised version received 04.03.24; accepted 16.03.24; published 05.06.24.

Please cite as:
Sawyer C, Carney R, Hassan L, Bucci S, Sainsbury J, Lovell K, Torous J, Firth J
Digital Lifestyle Interventions for Young People With Mental Illness: A Qualitative Study Among Mental Health Care Professionals
JMIR Hum Factors 2024;11:e53406
URL: https://humanfactors.jmir.org/2024/1/e53406
doi:10.2196/53406
PMID:38837191

©Chelsea Sawyer, Rebekah Carney, Lamiece Hassan, Sandra Bucci, John Sainsbury, Karina Lovell, John Torous, Joseph Firth. Originally published in JMIR Human Factors (https://humanfactors.jmir.org), 05.06.2024. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on https://humanfactors.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Viewpoint

Implementation of Anxiety UK's Ask Anxia Chatbot Service: Lessons Learned

Luke Collins[1], PhD; Niamh Nicholson[2], BSc; Nicky Lidbetter[2], PhD; Dave Smithson[2], HNC; Paul Baker[1], Prof Dr

[1]Linguistics and English Language, Lancaster University, Lancaster, United Kingdom

[2]Anxiety UK, Manchester, United Kingdom

Corresponding Author:
Luke Collins, PhD
Linguistics and English Language
Lancaster University
Economic and Social Research Council Centre for Corpus Approaches to Social Science
Bailrigg
Lancaster, LA1 4YW
United Kingdom
Phone: 44 1524 65201
Email: l.collins3@lancaster.ac.uk

Abstract

Chatbots are increasingly being applied in the context of health care, providing access to services when there are constraints on human resources. Simple, rule-based chatbots are suited to high-volume, repetitive tasks and can therefore be used effectively in providing users with important health information. In this Viewpoint paper, we report on the implementation of a chatbot service called Ask Anxia as part of a wider provision of information and support services offered by the UK national charity, Anxiety UK. We reflect on the changes made to the chatbot over the course of approximately 18 months as the Anxiety UK team monitored its performance and responded to recurrent themes in user queries by developing further information and services. We demonstrate how corpus linguistics can contribute to the evaluation of user queries and the optimization of responses. On the basis of these observations of how Anxiety UK has developed its own chatbot service, we offer recommendations for organizations looking to add automated conversational interfaces to their services.

(JMIR Hum Factors 2024;11:e53897) doi:10.2196/53897

KEYWORDS

chatbots; anxiety disorders; corpus linguistics; conversational agents; web-based care

Introduction

In the context of developing technologies, many businesses and services are turning to automated systems to provide users with information and accessible customer service. Among such tools, we find natural language processing systems, such as chatbots, that act as conversational interfaces, typically in lieu of interactions with human professionals. In health care, chatbots have a meaningful role to play, alongside other provisions, in increasing access to services, particularly in instances where there are restrictions in accessing face-to-face services [1,2]. Medical chatbots are already being used to provide and elicit information, create patient records, and discuss the results of clinical tests [3]. Furthermore, as Amiri and Karahanna [4] argue, health chatbots were shown to be particularly valuable in periods of quarantine as a response to the COVID-19 pandemic, in that "[t]heir scalability, wide accessibility, fast information dissemination, and substitution for in-person contact provide the functionality required to address the capacity expansion, social distancing requirements, and quick accurate information transmission needs of the public health response." Ultimately, chatbot tools and similar automated systems can make an important contribution to the provision of health information and support in the context of time and resource restraints.

There is a wide range of capabilities demonstrated in the deployment of conversational interfaces of varying complexity, from rule-based chatbots that produce prewritten responses based on recognizing programmed terms and phrases to embodied conversational agents manifesting as a computer-generated avatar and smart conversational interfaces such as Apple's Siri or Amazon's Alexa [1]. Nevertheless, simple conversational agents are increasingly used in executing tasks without the need for human involvement, including

XSL•FO
RenderX

booking appointments, purchasing merchandise, ordering food, and sharing information [5].

Research has shown that users respond positively toward the perceived convenience of medical chatbots, showing appreciation for swift information retrieval as an alternative to delays in scheduling a consultation, queuing in a phone service, or waiting for an email response [3]. There is wide acceptance of automated systems providing general health advice [2]. Furthermore, there are indications that computer services can reduce perceived stigma, in that users are more willing to disclose details about their health concerns to an automated system on the basis that they are regarded to be more trustworthy and nonjudgmental, reducing the potential for embarrassment [6,7]. In the mental health domain, chatbots and other kinds of conversational agents have been shown to assist in the diagnosis and reduction of symptoms among individuals with major depressive disorder, promote adherence, provide cognitive behavioral therapy, and cultivate a stronger therapeutic alliance compared with users' interactions with a clinician [7].

This study details the development and learning from the implementation of a chatbot service through the website of the mental health charity Anxiety UK, the largest national charity in the United Kingdom to offer support for anxiety disorders. The charity has created a chatbot service called Ask Anxia to complement its other support and information services. In this study, we summarize the patterns of queries submitted to the Ask Anxia service after approximately 18 months of its activation using procedures from corpus linguistics, which involves using software tools to compute frequency-based measures of naturally occurring language data [8]. In addition, we review the quality of Ask Anxia's responses based on manual coding. We offer some reflections on the development of the Ask Anxia service as "lessons learned," with the intention that these will be instructive to others seeking to incorporate a conversational agent into their provision of information and support.

Anxiety UK and Ask Anxia

Anxiety UK was established in 1970 and provides a wide range of support services and information for those affected by anxiety, stress, and anxiety-based depression. Anxiety disorders are characterized by excessive worry and fear [9] and are included among the "Common mental disorders" that are recorded as becoming increasingly prevalent in the United Kingdom [10]. The charity supports individuals from all over the United Kingdom and, in some cases, the rest of the world and has recently led on the development of an informal global alliance of not-for-profit anxiety organizations. Anxiety UK has a strong service delivery arm offering support via their helpline, therapy, peer support groups, and anxiety management courses. Most of, if not all, its volunteers, staff, and trustees have some experience of anxiety disorders. Anxiety UK states through all its communications, including the chatbot service, that it does not provide crisis support and directs those in need of such support to urgent care services such as the National Health Service and the charity, Samaritans.

Anxiety UK introduced an automated chatbot service, Ask Anxia, with the principal aim of offering an out-of-office-hours service to users, helping them to navigate more quickly to information that was already available, for example, through Anxiety UK's web pages. Furthermore, the Anxiety UK team found that a high number of user queries received by phone or email concerned administration issues, and so, providing such information through an automated chatbot was seen as a way to release staff members and helpline volunteers to attend to other responsibilities that demanded more critical and engaged attention, including providing real-time interactional support via the helpline. Anxia is now a registered trademark that includes but is not limited to computer and application software provided by Anxiety UK as part of their mental health services.

Ask Anxia is a simple, pattern-matching chatbot that has been programmed to recognize certain stimuli (specific terms or phrases) and generate a response, which has been composed by the Anxiety UK team. At the time of writing, Ask Anxia had a content bank of 315 unique responses that has been developed and refined since the service has been operational, and the Anxiety UK team continues to monitor these response options based on the range of queries that users submit. An overview of the categories of responses is provided in Table 1, indicating the types of terms that Ask Anxia has been programmed to recognize in user queries.

Ask Anxia was launched in the beginning of July 2021, and we have applied procedures from corpus analysis (discussed in the Developments section) to 56 weeks' worth of anonymized, aggregated user queries submitted to the service (up until the end of July 2022). This amounted to 14,359 queries consisting of 139,286 words.

XSL•FO
RenderX

Table 1. Recurring themes in queries to Ask Anxia and examples of terms used to determine a response.

Theme	Examples of pattern-matched terms
Request for help	help; support; advice; guidance
Looking for information on a specific anxiety type	GAD; health anxiety; OCD; PTSD; emetophobia; phobia
Physical symptoms	headache; chest pain; breathing; appetite; nausea; feel sick; dizzy
Psychological symptoms	intrusive thoughts; negative thinking; overthinking; constant worry
Information on a service	therapy; group; course; class; counselling; CBT; EMDR; resources
How to access a service	membership; cost; referral; book; sign up; join
Wanting to connect	talk; human; chat; agent
How to support others	family; partner; son; daughter; child; colleague
Getting involved with Anxiety UK	volunteering; approved therapist; fundraising; donate; placements
Diagnosis	do I have anxiety; diagnosis; symptoms
Medication	antidepressants; tablets; medication
Coping techniques	can't cope; what can I do; panic; now; thoughts; relax; sleep
Location	in person; areas; UK; Europe; face-to-face; online
Crisis	suicidal; self-harm; die

We recognize that there are important ethical considerations pertaining to the data, given that queries submitted to the Ask Anxia service are highly personal and relate to individuals' well-being. In the privacy notice that is posted on the Anxiety UK website, users are informed that interactions with Ask Anxia are reviewed as part of the procedures for improving the quality of the service and that these may be shared with third parties for the purposes of research. Participants are discouraged from including personal information in their queries, and any such information that appears in the original message has been redacted. To protect the personal experiences of those who have accessed the service, we have provided generic examples, where cited, to demonstrate the interactional dynamics between constructed user queries and Ask Anxia's (authentic) responses. Reported figures for word frequencies are based on original user queries. As part of a more recent update to the service (July 2022), a message encouraging users not to disclose personal information (such as name, address, and place of work) was added to the Ask Anxia header to ensure that this is visible. Furthermore, such information does not inform Ask Anxia's pattern-matching programming, and so, it will only hinder the identification of an appropriate response.

In addition to reporting commonly used terms and phrases, we refer to the quality coding carried out by the Anxiety UK team, which is explained in the next section. Our study, then, offers a critical evaluation of the contribution of the chatbot Ask Anxia to Anxiety UK's wider provision of services and helps us to understand the general patterns of what visitors to the site collectively seek, in terms of information and support. In the next section, we summarize the insights that we have gained through developing Ask Anxia's programming, as the service has evolved over time.

Developments

Overview

In this section, we summarize the developments that have been applied to the Ask Anxia service based on observations of user queries, including where potential misunderstandings in queries asked by users arose. We present these developments as the lessons we have learned through reviewing the various updates that have been applied to the service since its launch, which are likely to be informative to those looking to implement similar tools. The continued monitoring of the service has contributed to its optimization and generated insights into user expectations. The time stamps for user queries indicate that 57.2% (8213/14,359) of the queries were submitted outside of Anxiety UK's office hours (9:30 AM-5:30 PM), demonstrating that the Ask Anxia service is used when other contact services, such as the helpline, are closed. Indeed, one of the earliest modifications to Ask Anxia, in August 2021, was to remove the cap on how many queries it responded to, given its popularity.

Updates Based on Frequent Terms in User Queries

The pattern-matched terms presented in Table 1 were largely informed by the Anxiety UK team's own long-standing experiences of working with people seeking support for their, or a loved one's, experiences of anxiety. Of course, the queries submitted to Ask Anxia provide further indications of what users seek from the service. As such, alongside the Anxiety UK team's expert judgment, procedures from corpus linguistics can be drawn on to help identify topics and terms that are commonly cited by users, which can potentially highlight important areas for extending the existing information provision.

Corpus linguistics refers to a set or procedures for making quantitative and qualitative observations of the patterns of natural language use and can straightforwardly tell us, by way of a wordlist, for example, what the most common terms in our data are and how often they occur. We used the corpus analysis

XSL•FO
RenderX

tool *#LancsBox* [11] to examine the user queries. Researchers have found, however, that because of how the English language is structured, often the most common words largely remain the same across data sets (typically, *I*, *the*, *you*, *and*, *it*, etc). Indeed, the 5 most frequent terms occurring in the user queries of the Ask Anxia service, were *I*, *to*, *a*, *and*, and *hi*. As such, corpus linguists have developed the concept of keyness, enabling us to determine which words appear in our data more frequently, to a statistically significant degree, when compared with a corpus of larger or equal size [8]. A keyness analysis of the queries submitted to Ask Anxia through comparison with a 10 million–word corpus of general English spoken language, the British National Corpus 2014 [12], identified the keywords that are particularly characteristic of the language used by contributors in this context. The statistical measure used in this case was log likelihood, which established a confidence score indicating that the observed differences are not the result of chance. A threshold value of 15.13 was applied, which equates with a *P* value of <.001. The top 20 keywords are shown in Table 2 and ranked according to log likelihood value (not reported).

Table 2. Keywords in user queries to the Ask Anxia service (n=139,286).

Rank	Keyword	Frequency, n (%)
1	anxiety	2461 (1.77)
2	hi	2670 (1.92)
3	help	1588 (1.14)
4	hello	1332 (0.96)
5	yeah	13 (0.01)
6	am	1123 (0.81)
7	anxious	525 (0.4)
8	therapy	451 (0.3)
9	my	2358 (1.69)
10	panic	373 (0.3)
11	support	411 (0.3)
12	im	302 (0.2)
13	membership	310 (0.2)
14	struggling	334 (0.2)
15	i	7805 (5.6)
16	ok	259 (0.2)
17	how	1370 (0.98)
18	oh	47 (0.03)
19	feeling	394 (0.3)
20	can	1606 (1.15)

What is clear from the keywords is the topical focus on *anxiety* and the prevalence of appeals for *help* and *support* on the basis that users are *struggling*, *feelinganxious*, or experiencing *panic* (attacks), for instance. We can also see that queries are typically written in the first person (*I*, *my*, and *im*), take a question form (*how* and *can*), and have a relatively informal style (*hi*, *yeah*, *ok*, and *im*), that is, consistent with the instant messaging–like format through which users interact with Ask Anxia.

The prevalence of the terms *help* and queries about *therapy* and *membership* indicate that the Anxiety UK team had largely anticipated the themes most often captured in user queries, as indicated in Table 1. Nevertheless, the recurrence of particular terms, including at specific moments, has informed the continued refinement of Ask Anxia's responses and the information that is made available through the website. For example, in the week beginning September 13, 2021, keyness analysis showed that there was an increase in references to *fear* and *needle*, which coincided with booster doses of the COVID-19 vaccine being made available (to certain groups) and vaccines being approved for 12- to 15-year-olds, in anticipation of a new school term. Subsequently, the terms *fear* and *needle* appeared much more frequently in user queries. In response, Anxiety UK produced specific information concerning COVID-19 and related vaccines.

The Anxiety UK team has continued to extend Ask Anxia's response options since it was launched in July 2021. In addition, because of identifying themes arising from user enquires, the team has carried out the following activities:

- created factsheets specifically on perinatal anxiety, peri- and postmenopausal anxiety, and negative thoughts and catastrophizing
- created additional web content such as adding a do-it-yourself self-diagnosis section to the "About Anxiety"

XSL•FO
RenderX

page, extending the list of associated symptoms, explaining additional types of anxiety disorder such as dermatillomania, and adding further detail to the process of becoming a volunteer
- written and posted blogs on the topics of older people and anxiety, high functioning anxiety, highly sensitive people and anxiety, work anxiety, anxiety and appetite, autism and anxiety, anger and irritability with anxiety, returning to work post lockdown, and placements for students
- added entries to the frequently asked questions section relating to costs and arrangements for therapy
- extended member benefits, including researching the provision of fidget toys and fidget jewelry

As the updates have been informed by recurring user queries, we can expect that they will be of value to users generally, and by linking the updated information to Ask Anxia's responses, a greater number of queries can be addressed automatically, out of hours and without the need for human intervention. Being able to identify trends in information-seeking requests has enabled Anxiety UK to respond operationally and strategically to meet the needs of its beneficiaries.

Quality Coding

Each week, the Anxiety UK team manually coded a sample of Ask Anxia's responses to monitor quality, which we have labelled Good, Okay, Bad, or Puzzled. On average, the Anxiety UK team would code 155 queries per week (ranging between 0 and 408). Good responses provided the appropriate information based on the query and constituted the response option that the human coder would have selected. The following example shows how Ask Anxia responds to the mention of "social anxiety" and directs users to the appropriate information:

> *Details about social phobia/social anxiety can be found here* [link provided]

Responses coded as Okay were not necessarily the optimal response option but were still topically relevant. For instance, in the case of a user posting a query that indicated that they wanted to talk to someone about dealing with anxiety, a human reader is likely to recognize the importance of talking to *someone*, whereas a response from Ask Anxia, which would subsequently be coded as Okay, might respond to the mention of anxiety as follows:

> *We provide a wide range of services and information for those dealing with anxiety, stress, or anxiety-based depression. Check out our homepage as a start here: [link provided] to see our calendar of upcoming events and our latest news.*

In this instance, the user is still directed to information that is likely to be useful to them, even if this was not the primary purpose of their query. Where this points to a potential recalibration of the service is that Ask Anxia was already programmed with a response that more directly attends to the question of speaking with a (human) member of the team.

Bad responses appeared when there was misalignment with what the coders, and we, can perceive as the user's intended meaning, with Ask Anxia generating an irrelevant or inappropriate response when a more pertinent option was available. For example, following a query that mentioned "joining," that is, membership with Anxiety UK, Ask Anxia generated the following response pertaining to joining a webinar:

> *You can book on to our next webinar here:* [link provided]

Again, reviewing the cause of the misalignment highlights ways in which Ask Anxia can be improved.

Finally, certain responses are generated when a more specific alternative, relating to the topic or passage of interaction, is not available, such as:

> *Sorry, I am not sure how to answer that, do feel free to use our website search bar which may find the answer for you or contact our team directly; we are open Mon-Fri 09:30-17:30 (excluding bank holiday) and our contact list can be found here:* [link provided]

Such responses were coded as Puzzled.

Of the 14,359 queries, 8669 (60.37%) were subject to quality coding, and the distribution of these according to the different quality labels is shown in Table 3.

The quality coding figures provided in Table 3 indicate that Ask Anxia generally performed well, providing a Good response in two-thirds of cases. Furthermore, we can see how this coding was applied weekly, given that the Anxiety UK team made adjustments to Ask Anxia's response options based on what they observed in their coding. Figure 1 indicates how the queries were coded between October 2021 and December 2022, showing the proportion of Ask Anxia responses that received the codes Good, Okay, Bad, and Puzzled.

Table 3. Number of queries coded according to each quality code (N=14,359).

Code	Queries, n (%)
Good	5801 (66.92)
Okay	911 (10.51)
Bad	1537 (17.73)
Puzzled	420 (4.84)

XSL•FO
RenderX

Figure 1. Quality coding of Ask Anxia responses as a percentage during each week from September 27, 2021, to December 26, 2022.

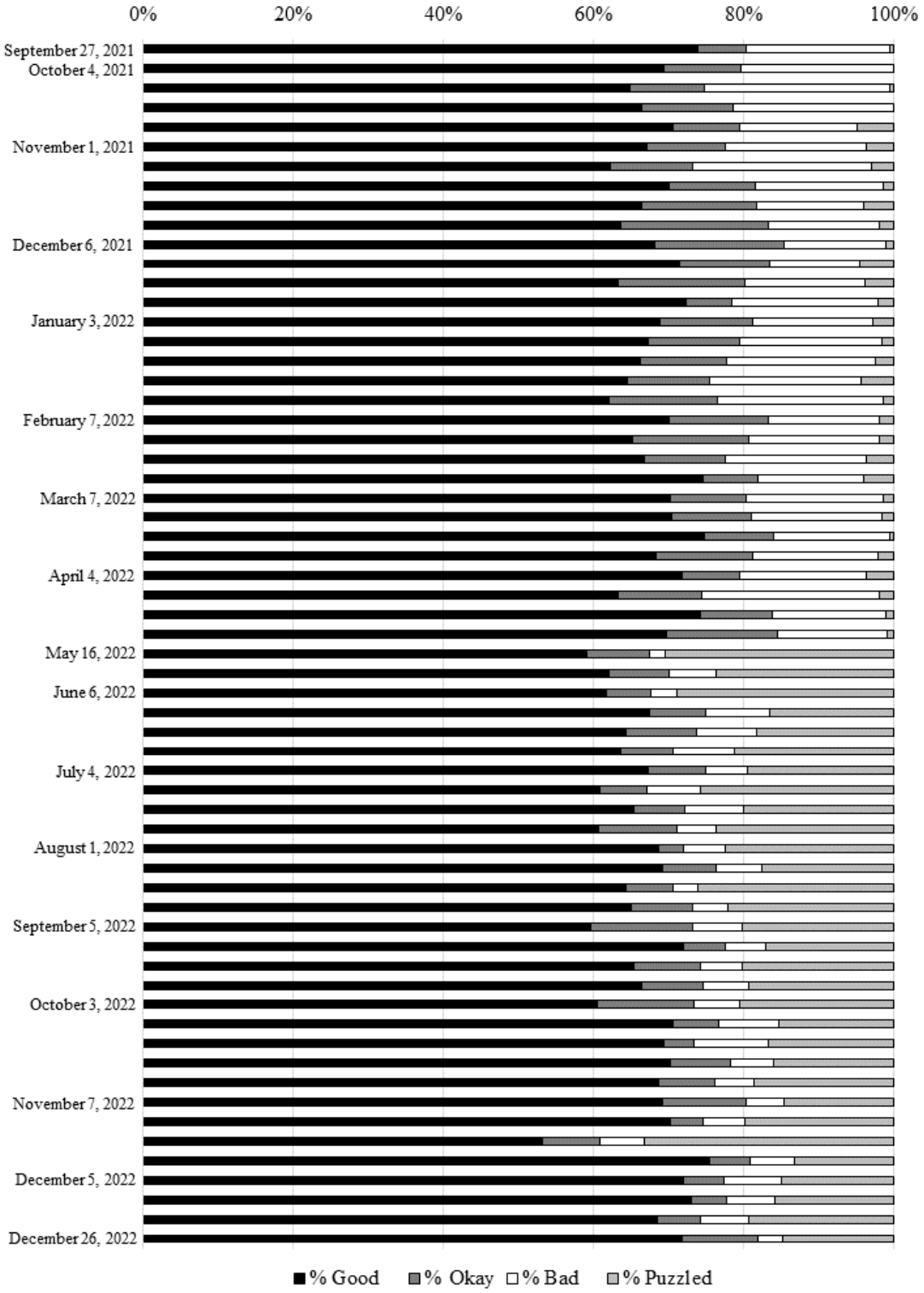


Figure 1 shows that the proportion of responses coded as Good was consistently >60% and that typically, Bad responses accounted for <30%. From May 2022 onward, there is a notable shift toward a smaller proportion of Bad and Okay responses and, instead, a greater number of Puzzled responses. As opposed to any change in the human evaluation of the responses or specific update to the software, this improvement reflects the ongoing work that the Anxiety UK team had been doing to calibrate the responses. This finding indicates that over time, Ask Anxia is less likely to generate an inappropriate or irrelevant response that could lead to disengagement from the user and more likely to provide a response that facilitates further engagement and opportunities for the user to find the relevant support.

The example of a Puzzled response provided earlier in this section on Quality coding was not initially one of the preprogrammed options but was added shortly after the launch (in September 2021), as the Anxiety UK team recognized a need

XSL•FO
RenderX

to indicate when an optimal response could not be offered. The necessity of a Puzzled response is to be expected, given that the Anxiety UK team is not reasonably going to be able to anticipate the full range of queries users could conceivably submit. Furthermore, in many cases, a Puzzled response is preferable to a Bad response because it encourages the participant to remain engaged and try again. In work analyzing approximately 20,000 conversational exchanges between customers and a task-oriented chatbot for a Taiwanese banking firm, Li et al [13] focused on the problem of "nonprogress" responses, where users abandoned the dialogue. They identified a number of "reformulation" strategies when progress was halted, including rephrasing; adding different words; repeating the same words; and to a lesser extent, removing words [13]. This suggests that prompting the user to reformulate their query or to try an alternative mode of engagement, which the Puzzled response does, is preferable to closing down the exchange. Often, users simplify their reformulated messages [14], which increases the probability for pattern matching and Ask Anxia finding a relevant response.

While a Puzzled response can be the appropriate response, for example, when there is no suitable prewritten response or information provision, monitoring the instances when such a response is elicited highlights areas where Anxiety UK can consider extending the response options or the information and services they provide through their website.

Pattern Matching

In this section, we report some of the modifications made to Ask Anxia designed to attend to features of user queries that can potentially disrupt the pattern-matching mechanism of simple automated chatbot systems. For instance, the Anxiety UK team became aware that the use of certain punctuation affected the ability of the bot to respond correctly to the query and duly updated the program to navigate around such characters.

The simplicity of a pattern-matching procedure is demonstrated when the input (the user query) is not identical to the stimulus the chatbot is programmed to recognize, which can occur with misspellings. In addition to informing us that the term *anxiety* appeared 2461 times in the user queries, the wordlist generated in *#LancsBox* also indicates that the following (likely) misspellings of "anxiety" occurred: *aniety, aniexty, aniexy, anixety, anixity, anixtey, anixty, anxciety, anxeity, anxety, anxiatey, anxiery,* and *anxiey.*

Recognizing common misspellings of relevant terms can help to minimize the number of cases in which the chatbot cannot identify an appropriate response, and while it may be unfeasible to program the service to recognize all possible variants, the wordlist allows us to identify the most common.

The use of negation can result in false negatives, in cases where users produce the relevant stimulus but deny or distance themselves from the concept in their proposition, for example, "not needle phobia." In such instances, while a chatbot can be programmed to recognize negation (in terms such as *not*, *isn't*, or *no*), the query does not provide the input to determine what is the impetus of the query, and so recognizing negation would not then help to identify a suitable response. In such cases, the onus may be on the user to deduce how the inappropriate response has been generated (ie, seeing the pattern matching with their original query) and to try reformulating their message. A more proactive response, on the part of the service provider, would be to program the chatbot to recognize negation and to generate a Puzzled-type response that prompts the user to reformulate their query.

Users' queries might also include additional pattern-matching terms that do not constitute the primary focus of their message but which nevertheless prompt a response This was often the case with longer, more complex query formulations in which multiple competing trigger terms appeared. In most cases, one of the terms would elicit a corresponding response, but this might simply be a greeting to a query that happened to begin with the word *hello*. AbuShawar and Atwell [15] compare a "first-word" approach to a "most significant word" approach with respect to programming chatbots; they explain that the "most significant" word is determined according to low frequency, on the basis that a low-frequency word is what distinguishes an utterance and will favor informational content over high-frequency function words, such as *a*, *to*, *in*, etc. This approach increases the probability that the tool is responding to a "topic" word rather than, say, a grammatical word; however, implementation as part of a simple, pattern-matching chatbot would require additional programming. In the case of Ask Anxia, the Anxiety UK team introduced a prompt in August 2022 that advised users to construct their queries in a simple and direct manner, thereby maximizing the potential for Ask Anxia to recognize a relevant term. Such a response can be generated on the basis of the length of the query (ie, character or line count).

Managing Expectations

In the previous section, we have seen that optimizing a chatbot service relies, to some extent, on the understanding of the user that, for example, simple direct queries are likely to produce the best results. As such, there is a degree of familiarity, or "literacy," that can help to ensure that users find the support and information that is of most benefit to them. Working toward this alignment between user goals and service is also a case of managing expectations, first and foremost in relation to what the Ask Anxia service is and can do.

A series of updates applied to Ask Anxia reflected the increasing explicitness with which the Anxiety UK team described the automated nature of the service. Shortly after its initial launch (July 2021), the team supplemented the initial "Hello" response with the following message:

> *Hello, I am Anxia the Anxiety UK chat bot. I am here to provide you with advice and information. A brief disclaimer: this is not a crisis service, if you feel you are at risk, please contract 111 or 999. Now, how may I help?*

Subsequently, Ask Anxia has been relabeled an "eHelper" (August 2021) to avoid potential stigma associated with "chatbots" [3], then later (February 2022), the introductory prompt was rephrased to read the following: "Ask Anxia—Not human but here to help." Reviewing user queries, it becomes

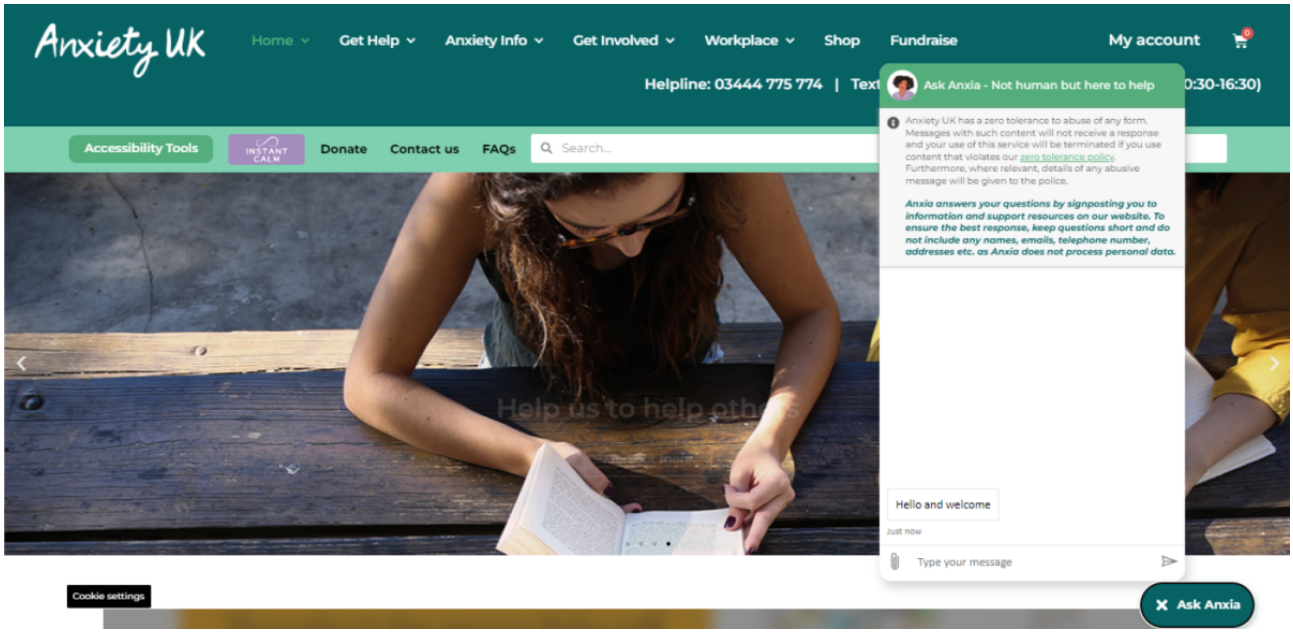
XSL•FO
RenderX

apparent why this clarification that the service is not operated by a human was required.

In Table 2, we say that one of the keywords for user queries was *oh*. Heritage [16] asserts that "where oh is produced as a response to information of some kind, it functions as a 'change of state' token; it registers, or at least enacts the registration of, a change in its producer's state of knowledge or information." In other words, the use of *oh* can indicate a degree of surprise or unexpectedness on the part of the recipient. When we refer to the queries, we see that often, this interjection reflected a realization on the part of the user that they were interacting with an automated service. The wordlist for the queries demonstrates the number of references to *bot* (87), *robot* (62), and *chatbot* (3), and we can extend our analysis in *#LancsBox* to determine frequencies of fixed phrases that include these terms, namely, *are you a robot* (21), *is this a bot* (16), *is this a robot* (14), *are you a bot* (12), etc.

On the one hand, this realization indicates a prior belief that the service was operated by a human and thereby might attest to the verisimilitude of the responses. On the other hand, the fact that this realization has come about indicates that such an illusion has been shattered, that is, because of an inappropriate response or perhaps because of repetition of the kind that is associated with pattern-matched chatbots with a limited number of responses [17]. Thus, with the aim of managing expectations and minimizing the potential for interactional trouble, the Anxiety UK team has worked toward more explicit signaling of the automated nature of the service (Figure 2). With this transparency, users can design their queries appropriately, and Anxiety UK can avoid too many instances where users become disillusioned by the potentially jarring realization that the interaction is not what they had presumed. Furthermore, we have established that some users may be more forthcoming knowing they are interacting with a nonhuman automated service [6].

Figure 2. The Ask Anxia chat window as it appears on the Anxiety UK website.



The Anxiety UK team had initially programmed Ask Anxia with a small number of light-hearted, conversational responses that contributed to a kind of persona, such as "I'm great, thanks for asking!" and "I don't have an age." Many of these responses were removed around September 2021 to October 2021, as they were often generated in inappropriate situations and there was a danger that they undermined the serious nature of the user query. Following their survey of motivations for using medical chatbots, Chang et al [3] determined that helping users acquire critical health information should take precedence over whether or not the chatbot appears empathetic or personable. Furthermore, when chatbots appear "humanlike," this can raise expectations about interactive capabilities, which in turn can negatively impact the interaction when the service's limitations are exposed [18]. Ultimately, transparency around the service's purpose and capabilities can help to avoid communicative misalignments and, given the high risk for responses that are designed to appear person like to actually appear flippant, such responses are arguably best avoided when user disengagement could potentially give rise to detrimental health consequences.

Finally, while Anxiety UK encourages users to be candid in their queries, on the basis that being direct will most likely mean that they can get the appropriate support, another series of developments to the Ask Anxia service has been to communicate a zero-tolerance policy for the use of profanities or abuse. A predefined response expressing this position was introduced in December 2021 and has since been subject to minor edits (also shown in Figure 2). Whether the use of profane language is motivated by frustration or is a more facetious "test" of the chatbot's capabilities (and there is evidence in the queries to indicate both), this does not include terms that are likely to be pattern matched to an informational resource. As such, there is no more preferable response for Ask Anxia to provide, other than to restate the zero-tolerance policy for such language, or alternatively, a response that encourages the user to reformulate their query. A summary of the updates described here is provided in Table 4.

XSL•FO
RenderX

Table 4. Summary of updates to the Ask Anxia service.

Date	Action
July 2021	• Extended welcome message clarifying that the service is operated by a bot as opposed to a human and to advise users that the service is not a crisis service.
August 2021	• Removed the limit cap on messages.
September 2021	• Removed a selection of prebuilt answers that were designed to make the bot seem more friendly and human-like but were giving inappropriate responses to sensitive queries. • Added a response advising users that the bot did not know how to answer their query (“Sorry I’m not sure how to answer that...”) as an alternative to generating poorly matched responses from the existing content bank.
December 2021	• Added a “zero-tolerance policy” response due to a minority of users using profanities. Further clarification of this response was enacted through minor amendments in January 2022 and also February 2022.
July 2022	• Added a response discouraging users from including personal data in their queries.
August 2022	• Added a response advising users to keep messages brief, on the basis that longer queries gave rise to confusion and poor responses from the chatbot.
October 2022	• Updated the content bank to recognize regular typing errors related to existing prompt terms.

Future Developments

It is worth highlighting some of the anticipated developments that will be implemented to continue to optimize the Ask Anxia service. These developments primarily orient around connecting users to the appropriate mode of service, for instance, providing the connection to contact a (human) operator when this is recognized in the user query. There are also instances in which a more informed response, beyond the level of detail provided in the preprogrammed replies, is required; in such cases, where the user query seems to rely upon more specific contextual or personal circumstances, Ask Anxia can direct users toward the helpline. The Anxiety UK team continues to refine the Puzzled responses to encourage further engagement from the user, for instance, providing the prompt, “Can you phrase this differently?” Finally, the Anxiety UK team is working on developing a mobile app that has the chatbot functionality embedded within it, thereby providing another arm of support and format to use the Ask Anxia service to reach a wider audience and attend to different user preferences.

Discussion

Organizations implementing pattern-matching chatbots for the purposes of providing information and support will benefit from continuous review of the response options and queries that users submit to their service. Furthermore, an initial set of programmed responses will likely need to be extended, and this will be informed by the nature of the queries that users submit. Our corpus analysis of frequently used terms in user queries to Ask Anxia demonstrated that the initial set of programmed responses was well aligned with the concerns of users but nevertheless helped to highlight areas where additional materials could prove to be useful. The manual quality coding of responses showed that Ask Anxia performs well, offering Good responses at a rate consistently >60%, and this procedure helped to identify areas where responses could be developed to address information gaps or otherwise refined to discern, for example, queries about needles generally and questions about specific vaccinations.

With respect to lessons learned through the implementation and review of the service, first, we have highlighted the informal nature of user queries, which often included ritualized greetings (*Hi* and *hello*). As such, it is useful to have a chatbot response that simply provides a greeting in kind. However, it is important to note that if a user greeting appears at the beginning of a more elaborate query, a response that attends to the topic of the query would be more appropriate.

Second, we have recommended that when an appropriate response cannot be readily identified, there is value in continuing the exchange, that is, encouraging the user to reformulate their query and thereby create additional input from which the chatbot can match an appropriate response. Researchers have highlighted the dangers of “nonprogress” responses that result in user disengagement [13]. Thus, while service providers are unlikely to be able to anticipate the full range of queries their users will submit, they can at least work to facilitate further engagement and use a preprogrammed, albeit uncertain reply to instruct participants on how best to elicit an acceptable response.

Third, we have seen that it is important to manage users’ expectations about what the tool can provide, which includes being explicit that the service is not provided by a human. Relatedly, responses that presented humanlike qualities proved to be of limited value, potentially raising expectations that the tool could offer humanlike judgments. Simple, pattern-matching chatbots such as Ask Anxia are best suited to “frequently asked questions”–type services, rather than more interactional, relationship-building tasks [17]. The benefit of these less-complex systems is that they are easier to program and implement and so can be adopted by service providers with minimal knowledge of the computational systems involved. It is important, nevertheless, to be cognizant of the limitations of such services. For instance, Ask Anxia does not track conversations over multiple turns but rather treats each post as

XSL•FO
RenderX

a new query; as such, any pertinent information provided at a previous turn is lost, and users may find themselves having to restate the fundamental purpose of their query. Similarly, the quality of Ask Anxia's performance is likely to diminish with longer, more complex queries, as it becomes more difficult to discern a singular, relevant prompt. Subsequently, users will be discouraged from providing contextual information (Figure 2) and are unlikely to receive personalized support in this mode. Simple chatbots, therefore, are arguably best used as part of an array of support options, including those which allow for more nuanced exchanges, for example, with a human provider over the telephone.

Laranjo et al [1] assert, based on a systematic review, that applications of chatbots in health care are in the early stages of development and evaluation. Furthermore, the systems used in health care lag behind those used in domains such as travel information and restaurant selection. As their deployment can have consequences for health outcomes, it is appropriate that such systems are continuously tested and evaluated. Language analysis is key to understanding both how users express themselves in queries to chatbots and the design of appropriate responses, and so, we advocate for the continued application of procedures such as those of corpus linguistics to support the extended use and performance of chatbots in health care.

Conclusions

The launch of the chatbot Ask Anxia was designed to support Anxiety UK in delivering information and support services to people concerned with anxiety disorders. The number of queries submitted to Ask Anxia, particularly out of hours, attests to the value of the service. In this study, we have demonstrated that procedures from corpus linguistics can help to identify patterns in user queries that reflect their needs and expectations of the service as well as direct us to where potential breakdowns in communication occur. For chatbot services to achieve optimal performance, human oversight is required, particularly during the first 6 to 12 months. Thereafter, less staff intervention is likely to be needed.

Acknowledgments

The authors would like to acknowledge the technical support provided by Roger Kadama in developing the Ask Anxia service. LC and PB were supported by funding from the Economic and Social Research Council (ES/R008906/1). The funding source had no role in the design of this study.

Authors' Contributions

All authors participated in the conceptualization of the paper and reviewed the manuscript. NN, NL, and DS carried out the manual coding of Ask Anxia's responses. LC and PB carried out the corpus analysis of user queries. LC wrote the manuscript with contributions from NN and NL on developments to the Ask Anxia service. NL conceptualized the Ask Anxia service. NN, NL, and DS contributed to the initial development of Ask Anxia and each continues to shape its development.

Conflicts of Interest

NN, NL, and DS are employed by the national charity, Anxiety UK, and were involved in the development of the Ask Anxia chatbot service.

References

1. Laranjo L, Dunn AG, Tong HL, Kocaballi AB, Chen J, Bashir R, et al. Conversational agents in healthcare: a systematic review. J Am Med Inform Assoc 2018 Sep 01;25(9):1248-1258 [FREE Full text] [doi: 10.1093/jamia/ocy072] [Medline: 30010941]
2. Nadarzynski T, Miles O, Cowie A, Ridge D. Acceptability of artificial intelligence (AI)-led chatbot services in healthcare: a mixed-methods study. Digit Health 2019 Aug 21;5:2055207619871808 [FREE Full text] [doi: 10.1177/2055207619871808] [Medline: 31467682]
3. Chang IC, Shih YS, Kuo KM. Why would you use medical chatbots? Interview and survey. Int J Med Inform 2022 Sep;165:104827. [doi: 10.1016/j.ijmedinf.2022.104827] [Medline: 35797921]
4. Amiri P, Karahanna E. Chatbot use cases in the COVID-19 public health response. J Am Med Inform Assoc 2022 Apr 13;29(5):1000-1010 [FREE Full text] [doi: 10.1093/jamia/ocac014] [Medline: 35137107]
5. Tudor Car L, Dhinagaran DA, Kyaw BM, Kowatsch T, Joty S, Theng YL, et al. Conversational agents in health care: scoping review and conceptual analysis. J Med Internet Res 2020 Aug 07;22(8):e17158 [FREE Full text] [doi: 10.2196/17158] [Medline: 32763886]
6. Palanica A, Flaschner P, Thommandram A, Li M, Fossat Y. Physicians' perceptions of chatbots in health care: cross-sectional web-based survey. J Med Internet Res 2019 Apr 05;21(4):e12887 [FREE Full text] [doi: 10.2196/12887] [Medline: 30950796]
7. Vaidyam AN, Wisniewski H, Halamka JD, Kashavan MS, Torous JB. Chatbots and conversational agents in mental health: a review of the psychiatric landscape. Can J Psychiatry 2019 Jul;64(7):456-464 [FREE Full text] [doi: 10.1177/0706743719828977] [Medline: 30897957]
8. Baker P, Hardie A, McEnery T. A Glossary of Corpus Linguistics. Edinburgh, UK: Edinburgh University Press; 2006.

XSL•FO
RenderX

9. American Psychiatric Association. Diagnostic and Statistical Manual of Mental Disorders. 5th edition. New York, NY: American Psychological Association; 2013.
10. McManus S, Bebbington P, Jenkins R, Brugha T. Mental health and wellbeing in England: adult psychiatric morbidity survey 2014. NHS Digital. 2016. URL: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/556596/apms-2014-full-rpt.pdf [accessed 2024-04-29]
11. Brezina V, Weill-Tessier P, McEnery A. #LancsBox v.6. Lancaster University. 2021. URL: http://corpora.lancs.ac.uk/lancsbox/ [accessed 2024-04-29]
12. Love R, Dembry C, Hardie A, Brezina V, McEnery T. The Spoken BNC2014: designing and building a spoken corpus of everyday conversations. Int J Corpus Linguist 2022 Jul 5;22(3):319-344. [doi: 10.1075/ijcl.22.3.02lov]
13. Li CH, Yeh SF, Chang TJ, Tsai MH, Chen K, Chang YJ. A conversation analysis of non-progress and coping strategies with a banking task-oriented chatbot. In: Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems. 2020 Presented at: CHI '20; April 25-30, 2020; Honolulu, HI p. 1-12 URL: https://dl.acm.org/doi/abs/10.1145/3313831.3376209 [doi: 10.1145/3313831.3376209]
14. Myers C, Furqan A, Nebolsky J, Caro K, Zhu J. Patterns for how users overcome obstacles in voice user interfaces. In: Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems. 2018 Presented at: CHI '18; April 21-26, 2018; Montreal, QC p. 1-7 URL: https://dl.acm.org/doi/10.1145/3173574.3173580 [doi: 10.1145/3173574.3173580]
15. AbuShawar B, Atwell E. ALICE chatbot: trials and outputs. Comp Y Sist 2015 Dec 27;19(4):625-632. [doi: 10.13053/cys-19-4-2326]
16. Heritage J. Oh-prefaced responses to inquiry. Lang Soc 1998 Jul 1;27(03):291-334. [doi: 10.1017/s0047404598003017]
17. Boucher EM, Harake NR, Ward HE, Stoeckl SE, Vargas J, Minkel J, et al. Artificially intelligent chatbots in digital mental health interventions: a review. Expert Rev Med Devices 2021 Dec 31;18(sup1):37-49 [FREE Full text] [doi: 10.1080/17434440.2021.2013200] [Medline: 34872429]
18. Gnewuch U, Morana S, Mädche A. Towards designing cooperative and social conversational agents for customer service. In: Proceedings of the 2017 International Conference on Information Systems - Transforming Society with Digital Innovation. 2017 Presented at: ICIS '17; December 10-13, 2017; Seoul, South Korea p. 1-13 URL: https://aisel.aisnet.org/icis2017/HCI/Presentations/1/ [doi: 10.17705/4icis]

Edited by A Kushniruk; submitted 23.10.23; peer-reviewed by M Chatzimina, AL MacNeill; comments to author 20.03.24; accepted 05.05.24; published 17.06.24.

Please cite as:
Collins L, Nicholson N, Lidbetter N, Smithson D, Baker P
Implementation of Anxiety UK's Ask Anxia Chatbot Service: Lessons Learned
JMIR Hum Factors 2024;11:e53897
URL: https://humanfactors.jmir.org/2024/1/e53897
doi:10.2196/53897
PMID:38885016

©Luke Collins, Niamh Nicholson, Nicky Lidbetter, Dave Smithson, Paul Baker. Originally published in JMIR Human Factors (https://humanfactors.jmir.org), 17.06.2024. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on https://humanfactors.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Alarm Management in Intensive Care: Qualitative Triangulation Study

Lina Mosch[1,2], MD; Meltem Sümer[3], MSc; Anne Rike Flint[1], MSc; Markus Feufel[3], MSc, PhD; Felix Balzer[1], MSc, MD, Dr rer nat; Frauke Mörike[3,4*], MA, PhD; Akira-Sebastian Poncette[1,2*], MD

[1]Institute of Medical Informatics, Charité – Universitätsmedizin Berlin, corporate member of Freie Universität Berlin and Humboldt Universität zu Berlin, Berlin, Germany

[2]Department of Anesthesiology and Intensive Care Medicine, Charité – Universitätsmedizin Berlin, corporate member of Freie Universität Berlin and Humboldt Universität zu Berlin, Berlin, Germany

[3]Division of Ergonomics, Department of Psychology and Ergonomics (IPA), Technische Universität Berlin, Berlin, Germany

[4]Department of Rehabilitation Sciences, Research Unit Work, Inclusion and Technology, Technische Universität Dortmund, Dortmund, Germany

[*]these authors contributed equally

Corresponding Author:
Lina Mosch, MD
Institute of Medical Informatics
Charité – Universitätsmedizin Berlin, corporate member of Freie Universität Berlin and Humboldt Universität zu Berlin
Charitéplatz 1
Berlin, 10117
Germany
Phone: 49 30450581 016
Email: lina.mosch@charite.de

Abstract

Background: The high number of unnecessary alarms in intensive care settings leads to alarm fatigue among staff and threatens patient safety. To develop and implement effective and sustainable solutions for alarm management in intensive care units (ICUs), an understanding of staff interactions with the patient monitoring system and alarm management practices is essential.

Objective: This study investigated the interaction of nurses and physicians with the patient monitoring system, their perceptions of alarm management, and smart alarm management solutions.

Methods: This explorative qualitative study with an ethnographic, multimethods approach was conducted in an ICU of a German university hospital. Using triangulation in data collection, 102 hours of field observations, 12 semistructured interviews with ICU staff members, and the results of a participatory task were analyzed. The data analysis followed an inductive, grounded theory approach.

Results: Nurses and physicians reported interacting with the continuous vital sign monitoring system for most of their work time and tasks. There were no established standards for alarm management; instead, nurses and physicians stated that alarms were addressed through ad hoc reactions, a practice they viewed as problematic. Staff members' perceptions of intelligent alarm management varied, but they highlighted the importance of understandable and traceable suggestions to increase trust and cognitive ease.

Conclusions: Staff members' interactions with the omnipresent patient monitoring system and its alarms are essential parts of ICU workflows and clinical decision-making. Alarm management standards and workflows have been shown to be deficient. Our observations, as well as staff feedback, suggest that changes are warranted. Solutions for alarm management should be designed and implemented with users, workflows, and real-world data at the core.

(JMIR Hum Factors 2024;11:e55571) doi:10.2196/55571

KEYWORDS

digital health; transdisciplinary research; technological innovation; patient-centered care; qualitative; ethnographic; ethnography; intensive care unit; ICU; intensive care; German; Germany; Europe; European; interview; interviews; alarm; alarms; intelligent; artificial intelligence; grounded theory; experience; experiences; attitude; attitudes; opinion; opinions; perception; perceptions; perspective; perspectives

XSL•FO
RenderX

Introduction

Background

In intensive care units (ICUs), continuous monitoring of multiple vital signs results in a high number of alarms [1] intended to inform staff of critical patient conditions, thus ensuring patient safety [2]. However, a large proportion of alarms do not require medical intervention [3], which, along with the sheer amount of alarms, disturb patient care [4]. In addition, the desensitization of ICU staff to alarms, "alarm fatigue" [5-7], is a threat to patient safety as it causes slower or no reactions to alarms [8,9]. Alarm fatigue is associated with working conditions and individual staff characteristics in deteriorating alarm monitoring performance [10]. Thus, for an improved monitoring performance, an understanding of workflows and the inner setting of a unit conducting vital sign monitoring is essential.

Alarm management standards have been proven effective in reducing unnecessary alarms [11,12]. Individual alarm thresholds, tailored to a patient based on previous monitoring data, alarm logs, and the medical data stored in the patient data management system (PDMS), would further improve alarm management [13,14]. One approach to this is using artificial intelligence (AI) to integrate these data and suggest alarm thresholds for an individual patient or actions to take after an alarm [15,16]. The research project *Intelligent Alarm Optimizer for the Intensive Care Unit* (INALO) follows this approach by creating a data set with semiautomatically annotated data (ie, alarms and the reactions to it) and using machine learning to predict the probability of an alarm to be actionable or nonactionable [17].

Before introducing new procedures or technologies in a complex sociotechnical environment such as an ICU, the characteristics of this setting and the individuals working there should be assessed [18,19], especially focusing on lived everyday work practices [20]. Since many alarm reduction solution approaches known in research have failed because of being too general [5], the interaction of the ICU staff with the patient monitoring system and the practiced alarm management should be investigated.

Objective

This qualitative study aimed:

1. to investigate the sociotechnical system ICU, with regard to the interaction of staff with the patient monitoring system and its alarms and
2. to understand the staff's perceptions of alarm management and the potential they see in the use of AI in this context.

We aimed to address the following research questions:

1. What is the sociotechnical role of patient monitoring in the work practice of nurses and physicians of the research site?
2. How does the staff, including nurses and physicians, of an ICU of a university hospital interact with the patient monitoring system, especially their interaction with alarms?
3. What is the attitude of physicians and nurses toward intelligent alarm management systems?

Methods

We consulted the Standards for Reporting Qualitative Research for reporting this research [21].

Qualitative Approach and Research Paradigm

This ethnographic study follows an explorative approach with (1) observation-based data collection triangulated with (2) semistructured interviews supported by (3) a self-reported overview of activities mapped out by each participant [22-24]. The research paradigm is postpositivist. We aimed to maintain a holistic, observing perspective while acknowledging that all the collected data and the derived findings are fallible, value laden, and need to be reflected from different researchers' perspectives [25]. The use of ethnographic methods fulfills the requirements of the postpositivist paradigm; it can create an understanding of social structures and capture the role and interaction with a technical system within those structures, revealing underlying sociotechnical relationships and patterns [26]. In order to build this comprehensive understanding, ethnographic research usually uses other methods in addition to observation [27]. Therefore, in addition to observation, semistructured interviews were conducted, and a participatory task was set for this work.

Researcher Characteristics and Reflexivity

The interdisciplinary research team consisted of a physician with work experience in medical informatics and anesthesiology and a research focus in implementation science and patient monitoring (LM); a psychologist enrolled in the master's program, *Human Factors*, exploring the interface between humans and technology (MS); a digital clinician scientist with work experience in medical informatics, anesthesiology, and internal medicine and a research focus on patient monitoring and alarm management (ASP); a professor of medical informatics (FB); a professor of inclusive work systems with research focus on ethnographic methods for technology design (FM); and a professor of ergonomics with a focus on human factors research in medical work environments (MF). MS was methodologically trained by FM and conducted the data collection from the etic perspective, that is, from the perspective of a person foreign to the field and the field of activity, supervised closely by FM throughout the field work. LM's and ASP's clinical perspectives provided a strong interdisciplinary contextualization of the data and informed the analysis. Throughout all phases of the research, FM, MS, and LM met regularly to discuss and reflect on the data collection process and preliminary findings. None of the research team members had a direct professional relationship with the research field.

Context

The motivation for this research, along with literature evidence of alarm fatigue, was a previous study that we had conducted to identify clinical requirements of future patient monitoring. We found that staff perceived alarm management as insufficient, threatening patient safety and disturbing workflows [28].

XSL•FO
RenderX

Setting

The research field of this study was an ICU of a German university hospital, where up to 21 patients can be treated in nine 2-bed rooms and six 1-bed (isolation) rooms. The devices used in the patient rooms were a continuous monitoring system (Philips IntelliVue) with sensors monitoring multiple vital parameters (eg, electrocardiogram, blood pressure, temperature, oxygen saturation, intracranial pressure, and electroencephalogram); a ventilator (Dräger Evita V800); infusion pumps (Agilia connect by Fresenius Kabi); and, if needed, a temperature management system (Arctic Sun by BD) or a dialyzer (various manufacturers). A table near the patient's bed with a computer enabled the caregivers to call up the PDMS to retrieve patient data and to document examination results. The PDMS used in this setting was COPRA 6 (COPRA System GmbH) [29], while the hospital information system used was i.s.h.med by Cerner [30]. The patient monitoring system installed on the ward included screens at the nurses' workstations directly outside the patient rooms and in the ward pulpit, the conference room, and a medication room. These screens displayed an overview of the vital signs of patients in adjacent rooms or of all patients in the ward.

Intelligent Alarm Optimizer Project

This study was conducted before the implementation and testing of INALO [31]. This project aims to improve alarm management for medical personnel using AI to decrease alarm loads, therefore easing the burden of alarm fatigue. The methodological approach included integrating and annotating data from multiple sources, including the hospital-wide patient monitoring system, the PDMS, and the hospital information system.

Access to the Field and Sampling Strategy

Access to the field was provided by the medical supervisor of this work (ASP) as well as the nursing management of the researched unit. The observations were announced to the entire ward staff by the ward manager via email, and the staff members were provided with the opportunity to address any concerns or queries they might have. A total of 12 observations were conducted by MS, each of which involved shadowing a nurse or physician for the duration of a shift (Table 1), for a total of 102 hours of observational data. The first 2 observations, shadowing both a nurse and a physician, were considered pilot observations. Their purpose was to become familiar with the environment of an ICU, test the methodology of observation and interviewing, and adapt the interview questions. The data from pilot observations were not used for later analysis.

Table 1. Number of observations divided by shift and profession and level of expertise, respectively[a].

Profession and level of experience	Early shift	Late shift	Night shift
Nurse (1.5-30 years of work experience in the ICU[b])	3	2	1
Resident physician (1-2 years of work experience in the ICU)	2	1	—[c]
Senior physician (7 years of work experience in the ICU)	1	—	—

[a]One participant was shadowed twice.

[b]ICU: intensive care unit.

[c]Not applicable.

The 5 nurses shadowed were suggested by the ward manager and scheduled with MS for different observation days after they had given written consent to participate in the study. Physicians were contacted by MS partly with the help of the medical supervisor or directly in the course of a field observation. In total, 4 physicians agreed to participate in the study.

The aim was to shadow a sample of staff members representative of the whole team, thus covering the greatest possible variety of perspectives in the research. Nurses and physicians of different ages and experience levels in an ICU were shadowed in all possible work shifts (early, late, and night shifts). The number of accompanying nurses and physicians per shift is listed in Table 1, where the experience level of working on an ICU is shown as the duration of work in intensive care.

Ethical Considerations

The study was approved by the ethics committee of the Charité Universitätsmedizin Berlin (EA4/218/20). Field notes and transcripts were anonymized or, if complete anonymization was not possible, pseudonymized. The data are stored on an internal institutional server, encrypted so that only the research team had access to it. Audio recordings were deleted after transcription. All staff members shadowed and interviewed for this study were informed about the research project and its aims and gave their written consent to participate. There was no compensation given to the staff members who were shadowed or interviewed.

Data Collection Methods and Instruments

Field Observation

The main data collection method in this study was field research in the form observation. In ethnographic research, the observation part varies by the degree of participation of the researcher in the field [32]. The setting of the field observations and the nonmedical background of the researcher suggest a nonparticipatory observer role (shadowing), in which events are observed in the background without involvement in any work processes. This methodology has been previously used in clinical settings [33-36]. Although the researcher observed the field from the background most of the time, she was also involved in conversations during the shifts by the shadowees. This gave her the opportunity to ask clarifying questions, which was beneficial in gaining insights. The observations were performed from April to October 2021. Breaks of up to 3 months

XSL•FO
RenderX

were taken for interim analysis of the data. This allowed for the readjustment of focus for the remaining observations and the evaluation of inductive thematic saturation [37], ultimately leading to the determination of the number of field observations as 7.

Semistructured Interviews

The interviews serve to explore the nonobservable topics relevant to the research question. They add to the observation method in the sense of methodological triangulation [22]. In particular, reactions to the planned AI-based INALO system and specific inquiries about the handling of the monitoring and its alarms were assessed. By not including the explicit term *AI* in the interview guide, the goal was to avoid potential biases or preconceived notions associated with the term, allowing for a more neutral and open interview discussion. *AI* was considered a polarized term, evoking personal opinions, as people may have subjective and varied views on the topic [38-40]. Consequently, the interviewees were provided with an explanation of the functionality of the INALO system, leaving out the specification *AI* (refer to question 5 in the interview guide in Textbox 1). The interview guide was developed through discussions of the research team and informed through the first 2 pilot observations

Textbox 1. Interview guide.

Questions

- Are there certain situations in which the sounding of an alarm is particularly inconvenient for you? Which ones?
- What role does managing alarms and especially their thresholds currently play during your work?
- How do you know it is time to adjust the alarm thresholds of certain parameters?
- When do you take the time to adjust an alarm threshold?
- Imagine a system, in addition to patient monitoring, that controls the alarm thresholds for oxygen saturation, arterial blood pressure and heart rate. If it registers that the current thresholds are no longer suitable for the patient in question, it suggests suitable alarm thresholds. What do you think about such a system?

In total, 7 interviews were performed individually with the 7 shadowed staff members (Table 1) after the end of a work shift in an open space to minimize outside distractions. The mean length of the interviews was 37:57 (range 56:22-15:05) minutes. We considered data saturation [37] including notes from field observations and interview transcripts. It was achieved after 6 field observations and, correspondingly, 6 interviews, which included early, late, and night shifts as well as nurses, resident, and senior physicians.

Participatory Task

In addition to the observation and the interview, following the triangulation strategy [41], a participatory task was conducted with the study participants. This helped gain insight into the representation of patient monitoring in the daily work routine and workflows of ICU staff members.

In a classic note-and-pen task, participants were asked to draw a pie chart of all the thoughts and tasks they think about during a work shift, with each piece of pie representing a thought, a task, or a task area. The size of the pie pieces was intended to show how the respondent's cognitive resources were divided during a shift. The participants were requested to prioritize the tasks that they had visualized based on their subjective perception. This was done to compare the estimated time allocated to each task in their workday with its perceived importance. The goal was note only to obtain indirect evidence on the importance of monitoring and alarms but also to understand the work routine in general from the respondent's perspective.

Data Collection Instruments and Technologies and Data Processing

Handwritten field notes were prepared following the methods of Emerson et al [42], and the date, location, and content of the observation were taken and promptly consolidated in MAXQDA (VERBI GmbH) [43] after each observation session. Interviews were recorded and transcribed according to the minimal transcript based on GAT 2 [44] and consolidated in MAXQDA. The participatory task was recorded in order to be able to use follow-up questions about the processing of the task as well as incidental comments for the later analysis.

Data Analysis

For inductive analysis, a grounded theory approach was followed [45]. The consolidated field notes and transcribed interview data from 6 observations were analyzed in MAXQDA 2022 and coded line by line in a first step, deriving thematic codes from the data while trying to neglect any relationship of the codes to the overarching research questions. In a second step, memoing was started to extract meaning from the data and to create and map an initial overview of themes [46]. This overview represented a preliminary code system after 6 observations. After all the gathered data were collected, coded, and tagged with open memos, the result consisted of 1336 coded data sections assigned to 47 parent codes and 70 thematically subordinate codes and tagged with 207 memos. Memos (ie, field notes with headlines) were organized thematically and considered in the context of their associated data sections (eg, interactions with the monitoring systems and alarm management). To ensure this thematic clustering according to the memo headline was rightful, we went back into the raw data (field notes) to evaluate whether the groups of memos with a similar headline could indeed be grouped together based on their raw data points. This helped to reevaluate and refine the derived themes from the memos (eg, integrating memos for more specific details and adding key points from field notes to memos from interviews and vice versa). When analyzing the data collected through the participatory task, the recorded topics

XSL•FO
RenderX

were written down, divided according to occupational group (nurses and physicians), and their occurrence was counted. As a result, an overview of superficial and seemingly less relevant task areas was created. The resulting themes from memos, codes, and topics from the participatory task were summarized and put in writing.

Techniques to Enhance Trustworthiness

Regular research meetings took place, where LM and MS discussed the findings and reflected on them. The code system (Multimedia Appendix 1) and memos were checked by both researchers under the supervision of FM. The combination of a psychologically trained Human Factors graduate student (MS) together with an expert in ethnographic research for work systems design (FM) and a physician with professional experience in an ICU and expertise in qualitative methods and implementation science (LM) was chosen to achieve the best balance of perspectives and topic prioritization. Interdisciplinary approaches are important to leverage the potential of research on the intersection of human-computer interaction, information systems, and health [47]. In addition, the multimethod study design with triangulation of complex data allowed for an increased credibility and trustworthiness of the results.

Results

Overview

The following 3 topics were identified from the data, following the research questions:

- the perception of the role of monitoring in the ICU
- the management and communication of vital sign limits (*dealing with alarms*)
- wishes and concerns regarding the intelligent alarm management system (eg, INALO)

An overview of nurses' and physicians' perceptions can be found in Table 2.

Table 2. Perceptions of monitoring and alarm management divided by professions.

Categories	Nurses	Physicians
The role of monitoring in the intensive care unit	• Use monitors at the bedside • Direct monitoring of vital signs at the screens	• Use monitors in the pulpit • Use nurses as monitoring filter
Dealing with alarms	• Subconscious adaptation to alarm patterns • First reaction to alarms • Implement alarm thresholds	• Subconscious adaptation to alarm patterns • Set and adjust alarm thresholds
Intelligent alarm management	• Concerned about the system's functionality • Highlighted the importance of understanding the system's operational principles • Concerned about excessive confidence	• Perceived as positive for lowering alarm burden • Perceived as valuable for new patients in the ward or upon returning from leave

The Perception of the Role of Monitoring in the ICU

Continuous monitoring of vital signs is an essential component of intensive care management of critically ill patients and a ubiquitous part of the ICU under study. According to a senior physician, even a short period without monitoring would be "grossly negligent."

The nursing staff primarily used the monitors located directly at the patient's bedside (ie, *direct* monitoring), as well as the overview screens at their workstations outside the patient rooms. During lengthy nursing activities in a room, it was observed that nurses used the monitoring system's function to display other patients' vital sign data on the screen of the bed where they were busy. Due to the larger number of patients under their care, physicians relied on the feedback from the nursing staff in the event of a critical situation (nurses as *monitoring filter*). They also paid closer attention to alarms on the overview monitors at the pulpit, and it was important that alarm thresholds were well set so that potentially life-threatening situations did not go unnoticed for long:

> *Especially when patients deteriorate, you notice that very often because the alarms are triggered all the time...These alarms...also signal something to us[...], without us seeing the patient directly[...], and if [the thresholds] were always set lower and it simply no longer alerts me, I'd only notice it when it's really bad or too bad or the catastrophe is there, and I wouldn't want that.* [Interview, physician]

When making therapy decisions, physicians considered the target values for specific parameters that had to be adhered to, based on therapy plans. Continuous monitoring of vital signs was crucial for physicians in ensuring that these target values are met.

Although monitoring was used in various areas of activity, it was not considered a central activity in and of itself. In the participatory task, monitoring was not mentioned by anyone. Alarms occurred in 2 of the 5 participatory tasks completed. However, when asked to indicate the role of monitoring in all the tasks written down, interviewees assigned monitoring to most of them (Multimedia Appendix 2).

Dealing With Alarms

The predominant observation regarding the handling of alarms was that they were directly paused or completely ignored, unless it was a red alarm. Often, there was no further reaction or intervention following an alarm.

However, as described in the previous section on monitoring, even an alarm that did not require intervention could provide helpful information about the progress of an individual patient's

XSL•FO
RenderX

condition. Both physicians and nurses showed a reliance on the alarms' feedback to be passively informed of the change in their patients' condition (subconscious adaptation to alarm patterns). In contrast, the alarms interfered with their work:

> *[...]as you may have noticed, and if this alarm comes up all the time...I actually look at it [the overview monitor] a lot. That disturbs me in my regular workflow.* [Interview, physician]

The alarm thresholds of vital parameters were set and adjusted either based on a patient's admission to the ward or return from the operating room or based on evaluations of the thresholds during the course of therapy. In the case of newly admitted patients, the threshold values were determined directly on admission, with the parameters being set according to standard values or having to be determined individually depending on the clinical picture. Nurses, in consultation with physicians, adjusted alarm thresholds according to the patients' condition to avoid unnecessary alarms. The decision to adjust thresholds varied based on the nurse's experience and familiarity with the patient. However, these informal procedures sometimes led to inappropriate thresholds on the next shift, as they were not always updated in the COPRA PDMS [29]:

> *...I come to the ward, and I don't know the patient. COPRA is binding, what is documented there should be correct and implemented at the bedside. If you're lucky, it says everything about which blood pressure values are desired...You have to ask about all that [if it is not entered correctly] and that's sometimes a bit annoying, until you find out which doctor is responsible for your room, where he is and what his telephone number is. It doesn't have to be like that.* [Interview, nurse]

Ideally, the determined threshold values were reevaluated daily in the morning shift. Often this was not the case due to a lack of time. Rather, this reevaluation only happened, if a new test result or diagnosis was found or in the case of disturbing, conspicuous, and unnecessary alarms. The chain of communication regarding the change of alarm thresholds from their determination to their implementation in monitoring ran hierarchically downwards, from senior physicians to assistant physicians to nursing.

If the alarm of a certain parameter would continue despite intervening measures or if no measures had been necessary at all, the alarm thresholds would be adjusted. Physicians and nurses frequently sought each other out for direct and immediate instructions. The official way of issuing instructions was via the COPRA PDMS, but according to staff, this was often too time-consuming, with delays for actual implementation:

> *And this adjustment [of the thresholds]...doesn't take place so much on this piece of paper, but rather on the patient himself, that I say, ok, we have to change the threshold here[...], then you discuss it next to the device itself with the nursing staff.* [Interview, physician]

Staff Perceptions of Intelligent Alarm Management

There were mixed reactions regarding an automated system to improve alarm management on the part of the nursing staff interviewed. Some expressed reservations, while others were open to the system described but said it would have to work well. There were concerns about the functionality of an intelligent system. The importance of understanding how such a system operates was highlighted. Staff members argued that currently a lot of responsibility was being handed over to technology anyway and that it would be more appropriate to be cautious about overconfidence:

> *I probably rely more on my own experience and won't put it in the hands of a machine that I don't know how it's programmed, what kind of things it reacts to, or what it takes as a basis for the recommendations. That would make a lot of sense for young colleagues who don't have a lot of experience yet, but you end up relying more on your own experience.* [Interview, nurse]

> *The patients, the clinical pictures, and that's all so individual and different that I can't imagine that a machine can [suggest thresholds].* [Interview, nurse]

Second, the statement was made that with clinical experience, one would notice when a threshold should be changed, and therefore, no system would be needed to do so:

> *I know already by my experience that I must intervene starting from certain values...you observe the patient, you observe the monitor and you already see...that you have to do something.* [Interview, nurse]

Thus, staff members saw the added value of an intelligent alarm system for less experienced colleagues and for leasing nurses, as the former would not yet have developed a sense of threshold assessment and the latter would not be familiar with patients. In addition, staff members saw a benefit in the use of the system to ensure regular evaluation and adjustment of alarm thresholds, preventing them from being forgotten.

Physicians reported that any solutions leading to a reduction in alarms would be welcome due to frequent alarms disturbing their concentration during work, especially in the ward pulpit:

> *[...]If it leads to the fact that the alarms are optimized, then I think it is already a relief, because I get fewer false alarms and thus, I can do the rest of the work more concentrated. So, that would already be an advantage. Certainly, it doesn't take away my cognitive work whether these alarms are, so to speak, suitable or not, I have to do that, but that's also my job.* [Interview, physician]

Finally, the advantage of such a system was observed by physicians for patients who were new to the ward and whose condition they did not yet know. The physicians said that it was also helpful to receive threshold suggestions when returning to the ward after a leave of absence, as they also needed to reacquaint themselves with patients.

XSL•FO
RenderX

Discussion

Principal Findings

Patient monitoring is an integral part of the work routine in the ICU. However, standards for working with the system were not implemented in clinical routine in the studied ICU. Alarm management is one of the core interactions with the monitoring system. Most alarms were confirmed without a reaction or intervention. The setting and adjusting of alarm thresholds were performed upon the arrival of a patient in the ICU or over the course of treatment by (senior) physicians and implemented by nursing staff. Perceptions of an intelligent system to suggest alarm thresholds varied: physicians saw potential advantages in a relief of the flood of (nonactionable) alarms by individualized alarm thresholds. In contrast, both nurses and physicians were skeptical about the capability of an automated system to perform the complex task of interpreting alarms and suggesting thresholds, encompassing the integration of different data and information. The importance of knowing how such a system worked internally and how it took decisions was highlighted.

Guidelines for Patient Monitoring and Alarm Management in the ICU

This work highlights the essential role of vital sign monitoring in the daily routines of all health care professionals in the ICU. Yet, to the authors' knowledge, interprofessional, systematic standard, and evidence-based guidelines for patient monitoring in clinical practice remain limited. Alarm management, a fundamental aspect of interacting with monitoring systems, warrants special attention, since alarm overload in ICUs hampers adequate response by personnel [48]. The American Association of Critical Care Nurses published recommendations for clinical alarm management, focusing on personalization of alarms and interprofessional alarm management strategies and highlighting the potential of smart alarms [49,50]. Our research supports the need for a wider clinical implementation of such recommendations, as we saw that a majority of alarms are disregarded or confirmed without any medical response or intervention [51,52].

Take Aways for Future Alarm Management

It is crucial to understand user perspectives when developing intelligent (AI-based) systems for alarm management in intensive care medicine [49]. The following aspects should be considered.

- Skepticism exists among staff members regarding the ability of an intelligent system to integrate diverse data formats and information to effectively interpret alarms and propose suitable thresholds. It might stem from a lack of knowledge and understanding of the internal workflows and decision-making processes of such systems. The potential benefits of integrating intelligent systems to automatically suggest personalized alarm thresholds and alleviate alarm fatigue [16,53] were acknowledged by physicians in our study.
- Standardized workflows for alarm management were not existent, and alarm thresholds and their adaptation were communicated irregularly and in a variety of ways. Well-defined and clearly communicated standard operating procedures (SOPs) for alarm management could address some of the challenges faced by health care professionals in the ICU [54,55]. Intelligent alarm management solutions should be implemented in clinical environment with well-established alarm management SOPs.
- Alarm management was performed mainly after new test results for the patients were received, a diagnosis was confirmed, or when the alarms were frequent and obviously unnecessary. This indicates that the data used for future alarm management systems based on AI need to mirror highly individual and complex medical conditions. Patients and clinical routines can differ even for various wards in the same clinic.

Extrapolating from these findings, we advise the following for ICU AI projects.

- For a truly effective implementation of AI systems, the ICU staff must be integrated in the design and implementation process, as well as possess adequate AI literacy. Encouraging and providing training to understand and use AI can empower ICU staff to embrace AI technologies confidently.
- As we need standards for alarm management workflows today, standardization is all the more essential in the context of integrating intelligent alarm management (AI-based) technologies. Significant workflow changes could be evoked and need to be considered by ICU leaders and implementation managers.

Limitations

In this study, the research focus lay on a single setting (ICU), which was suggested as a methodological approach by Wilken et al [5], allowing for an in-depth analysis of the contributing factors to an ICU's alarm management and dealing with the patient monitoring system. In addition to the stated benefits of this approach by the authors, it comes with limitations. On the one hand, the focus on a single ICU's cultural peculiarities restricts the transferability of the derived recommendations to other settings. The results and methods may serve, however, as inspiration and study protocol for evaluating an ICU's alarm management. By contrast, following a qualitative research approach, the number of shadowees and the number of shadowing days had to fit in a reasonable time schedule for the research conduction. The unit manager's suggestion of shadowees may have introduced a selection bias, despite being informed about the goal of diverse perspective shadowing. The results presented here are therefore only an excerpt of the reality in the ICU, which must be taken into account when interpreting them.

Conclusions

Our study highlights that interactions with the patient monitoring system and its alarms are a core part of tasks and workflows in the ICU. Alarm management tasks are performed based on ad hoc responses to clinical events; responsibilities are not well defined, and there is no standardized workflow or an SOP. Staff members were not satisfied with the current alarm management,

XSL•FO
RenderX

which emphasizes a need for standard and clinician-centered guidelines in this field. Establishing SOPs for configuring and responding to alarms and considering local patient and workflow characteristics can streamline tasks and enhance the overall efficiency of care delivery. Systems that enable an intelligent alarm management to reduce alarm fatigue among staff members should be designed to make understandable and traceable suggestions, while health care professionals should be empowered to use them meaningfully through digital health literacy. By establishing these standards and thoughtfully incorporating AI into clinical workflows, health care institutions could enhance patient safety and relieve staff and patients from alarm-induced stress. To explore this effect on outcomes, more research in this field is needed.

Acknowledgments

The authors express their gratitude to the volunteers and the ward staff members. The project *Intelligent Alarm Optimizer for the Intensive Care Unit* (INALO) is funded by the German Federal Ministry of Education and Research under grant 16SV8559. We acknowledge financial support from the Open-Access Publication Fund of Charité Universitätsmedizin Berlin and the German Research Foundation (DFG). We used the generative AI tool, ChatGPT by OpenAI, after the manuscript was written to rephrase single (approximately 12) sentences, to improve readability. The original ChatGPT transcripts can be obtained upon request from the first author (LM).

ASP and ARF received funding by the German Federal Ministry of Education and Research under grant 16SV8559 as part of the project INALO. The authors acknowledge financial support from the Open-Access Publication Fund of Charité Universitätsmedizin Berlin and the German Research Foundation (DFG).

Data Availability

The data sets generated and analyzed during this study are not publicly available because of data privacy; however, they are available from the corresponding author (LM) upon reasonable request.

Authors' Contributions

The study was conceived by ASP, LM, FM, and MS. FM designed the ethnographic study. MS conducted the data acquisition, supported by LM and ASP and under the methodological supervision of FM. Data analysis was performed by MS, LM, and FM. LM, MS, and ARF wrote the manuscript. ASP, FB, and MF supervised all parts of the study. All authors critically reviewed and approved the manuscript.

Conflicts of Interest

ASP and ARF received funding by the German Federal Ministry of Education and Research under grant 16SV8559 as part of the project INALO.

Multimedia Appendix 1
Codes and subcodes derived during data analysis from field notes, interview documentation, and documentation of the participatory task.
[DOCX File , 31 KB - humanfactors_v11i1e55571_app1.docx]

Multimedia Appendix 2
Indicated activities of a physician and their perceived proportion of the workday adapted from the results of the participatory task performed by a physician shadowee. The tasks with monitoring influence are indicated with the monitor icon.
[PNG File , 173 KB - humanfactors_v11i1e55571_app2.png]

References

1. Jones K. Alarm fatigue a top patient safety hazard. CMAJ 2014 Feb 18;186(3):178 [FREE Full text] [doi: 10.1503/cmaj.109-4696] [Medline: 24418978]
2. Moreno RP, Rhodes A, Donchin Y, European Society of Intensive Care. Patient safety in intensive care medicine: the Declaration of Vienna. Intensive Care Med 2009 Oct;35(10):1667-1672 [FREE Full text] [doi: 10.1007/s00134-009-1621-2] [Medline: 19697007]
3. Honan L, Funk M, Maynard M, Fahs D, Clark JT, David Y. Nurses' perspectives on clinical alarms. Am J Crit Care 2015 Sep;24(5):387-395. [doi: 10.4037/ajcc2015552] [Medline: 26330431]
4. Poncette AS, Mosch L, Spies C, Schmieding M, Schiefenhövel F, Krampe H, et al. Improvements in patient monitoring in the intensive care unit: survey study. J Med Internet Res 2020 Jun 19;22(6):e19091 [FREE Full text] [doi: 10.2196/19091] [Medline: 32459655]

XSL•FO
RenderX

5. Wilken M, Hüske-Kraus D, Klausen A, Koch C, Schlauch W, Röhrig R. Alarm fatigue: causes and effects. In: Proceedings of the 62nd Annual Meeting of the German Association of Medical Informatics, Biometry and Epidemiology. 2017 Presented at: GMDS '17; August 29, 2017; Oldenburg, Germany p. 107-111 URL: https://ebooks.iospress.nl/publication/47517 [doi: 10.3233/978-1-61499-808-2-107]
6. Sendelbach S. Alarm fatigue. Nurs Clin North Am 2012 Sep;47(3):375-382. [doi: 10.1016/j.cnur.2012.05.009] [Medline: 22920428]
7. Despins LA. Factors influencing when intensive care unit nurses go to the bedside to investigate patient related alarms: a descriptive qualitative study. Intensive Crit Care Nurs 2017 Dec;43:101-107. [doi: 10.1016/j.iccn.2017.04.003] [Medline: 28552259]
8. Paine CW, Goel VV, Ely E, Stave CD, Stemler S, Zander M, et al. Systematic review of physiologic monitor alarm characteristics and pragmatic interventions to reduce alarm frequency. J Hosp Med 2016 Feb;11(2):136-144 [FREE Full text] [doi: 10.1002/jhm.2520] [Medline: 26663904]
9. Joint Commission. Medical device alarm safety in hospitals. Sentinel Event Alert 2013 Apr 08(50):1-3. [Medline: 23767076]
10. Deb S, Claudio D. Alarm fatigue and its influence on staff performance. IIE Trans Healthc Syst Eng 2015 Jul 29;5(3):183-196. [doi: 10.1080/19488300.2015.1062065]
11. Hüske-Kraus D, Wilken M, Röhrig R. Measuring alarm system quality in intensive care units (work in progress). EasyChair. Preprint posted online May 6, 2018 [FREE Full text] [doi: 10.29007/5vpj]
12. Bach TA, Berglund LM, Turk E. Managing alarm systems for quality and safety in the hospital setting. BMJ Open Qual 2018 Jul 25;7(3):e000202 [FREE Full text] [doi: 10.1136/bmjoq-2017-000202] [Medline: 30094341]
13. Cosper P, Zellinger M, Enebo A, Jacques S, Razzano L, Flack MN. Improving clinical alarm management: guidance and strategies. Biomed Instrum Technol 2017;51(2):109-115. [doi: 10.2345/0899-8205-51.2.109] [Medline: 28296432]
14. Balzer F, Agha-Mir-Salim L, Ziemert N, Schmieding M, Mosch L, Prendke M, et al. Staff perspectives on the influence of patient characteristics on alarm management in the intensive care unit: a cross-sectional survey study. BMC Health Serv Res 2023 Jul 05;23(1):729 [FREE Full text] [doi: 10.1186/s12913-023-09688-x] [Medline: 37407989]
15. Huo J, Wung S, Roveda J, Li A. Reducing false alarms in intensive care units: a scoping review. Explor Res Hypothesis Med 2022 Jun 2;8(1):57-64. [doi: 10.14218/erhm.2022.00026]
16. Chromik J, Klopfenstein SA, Pfitzner B, Sinno ZC, Arnrich B, Balzer F, et al. Computational approaches to alleviate alarm fatigue in intensive care medicine: a systematic literature review. Front Digit Health 2022;4:843747 [FREE Full text] [doi: 10.3389/fdgth.2022.843747] [Medline: 36052315]
17. INALO: Intelligenter Alarmoptimierer für die Intensivstation. Miteinander durch Innovation. URL: https://www.interaktive-technologien.de/projekte/inalo [accessed 2023-12-04]
18. Powell BJ, Waltz TJ, Chinman MJ, Damschroder LJ, Smith JL, Matthieu MM, et al. A refined compilation of implementation strategies: results from the Expert Recommendations for Implementing Change (ERIC) project. Implement Sci 2015 Feb 12;10(1):21 [FREE Full text] [doi: 10.1186/s13012-015-0209-1] [Medline: 25889199]
19. Mosch LK, Poncette AS, Spies C, Weber-Carstens S, Schieler M, Krampe H, et al. Creation of an evidence-based implementation framework for digital health technology in the intensive care unit: qualitative study. JMIR Form Res 2022 Apr 08;6(4):e22866 [FREE Full text] [doi: 10.2196/22866] [Medline: 35394445]
20. Mörike F. Inverted hierarchies on the shop floor: the organisational layer of workarounds for collaboration in the metal industry. Comput Supp Coop Work 2021 Dec 02;31(1):111-147. [doi: 10.1007/S10606-021-09415-2]
21. O'Brien BC, Harris IB, Beckman TJ, Reed DA, Cook DA. Standards for reporting qualitative research: a synthesis of recommendations. Acad Med 2014 Sep;89(9):1245-1251 [FREE Full text] [doi: 10.1097/ACM.0000000000000388] [Medline: 24979285]
22. Flick U. Triangulation in der Ethnographie. In: Flick U, editor. Triangulation: Eine Einführung. Cham, Switzerland: Springer; 2011:51-74.
23. Patton MQ. Qualitative Research & Evaluation Methods: Integrating Theory and Practice. Thousand Oaks, CA: SAGE Publications; 2014.
24. Vocke M, Mörike F. E-Mails and teamrooms: immaterielle arbeit und digitale kommunikationstechnologien bei wissensarbeiter*innen in multinationalen unternehmen. In: Proceedings of the 66th Congress of the Society for Ergonomics on Digital Work, Digital Change, Digital People. 2020 Presented at: DWCP '20; March 16-18, 2020; Berlin, Germany URL: https://www.researchgate.net/publication/340005897_Chats_E-Mails_Teamrooms_Immaterielle_Arbeit_und_digitale_Kommunikationstechnologien_bei_Wissensarbeiterinnen_in_multinationalen_Unternehmen
25. Kankam PK. The use of paradigms in information research. Libr Inform Sci Res 2019 Apr;41(2):85-92. [doi: 10.1016/j.lisr.2019.04.003]
26. Dourish P. Implications for design. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. 2006 Presented at: CHI '06; April 22-27, 2006; Montréal, QC p. 541-550 URL: https://dl.acm.org/doi/10.1145/1124772.1124855 [doi: 10.1145/1124772.1124855]
27. Hammersley M. Reading Ethnographic Research. New York, NY: Routledge; 2016.

XSL•FO
RenderX

28. Poncette AS, Spies C, Mosch L, Schieler M, Weber-Carstens S, Krampe H, et al. Clinical requirements of future patient monitoring in the intensive care unit: qualitative study. JMIR Med Inform 2019 Apr 30;7(2):e13064 [FREE Full text] [doi: 10.2196/13064] [Medline: 31038467]
29. COPRA Patient Data Management Systems (PDMS). COPRA System GmbH. URL: https://copra-system.de [accessed 2023-12-16]
30. i.s.h.med: das universelle, in die SAP-welt integrierte KIS. Oracle Cerner. URL: https://www.cerner.com/de/de/loesungen/ishmed [accessed 2023-12-16]
31. Home page. INALO Intelligenter Alarmoptimierer für die Intensivstation. URL: https://www.inalo.ai [accessed 2023-02-25]
32. Czarniawska-Joerges B. Shadowing: And Other Techniques for Doing Fieldwork in Modern Societies. Copenhagen, Denmark: Copenhagen Business School Press; 2012.
33. Mörike F, Spiehl HL, Feufel MA. Workarounds in the shadow system: an ethnographic study of requirements for documentation and cooperation in a clinical advisory center. Hum Factors 2024 Mar;66(3):636-646. [doi: 10.1177/00187208221087013] [Medline: 35471848]
34. Leslie M, Paradis E, Gropper MA, Kitto S, Reeves S, Pronovost P. An ethnographic study of health information technology use in three intensive care units. Health Serv Res 2017 Aug 25;52(4):1330-1348 [FREE Full text] [doi: 10.1111/1475-6773.12466] [Medline: 28124443]
35. Smith A, Goodwin D, Mort M, Pope C. Expertise in practice: an ethnographic study exploring acquisition and use of knowledge in anaesthesia. Br J Anaesth 2003 Sep;91(3):319-328 [FREE Full text] [doi: 10.1093/bja/aeg180] [Medline: 12925468]
36. Schmidt C, Mörike F, Feufel MA. Ein digitaler Dritter im Ärzt*in-Patient*in-Gespräch: wie ein online-gestütztes tool beratungen standardisiert und personalisiert. Curare 2021;45(2022):77-100 [FREE Full text]
37. Saunders B, Sim J, Kingstone T, Baker S, Waterfield J, Bartlam B, et al. Saturation in qualitative research: exploring its conceptualization and operationalization. Qual Quant 2018;52(4):1893-1907 [FREE Full text] [doi: 10.1007/s11135-017-0574-8] [Medline: 29937585]
38. Fast E, Horvitz E. Long-term trends in the public perception of artificial intelligence. Proc AAAI Conf Artif Intell 2017 Feb 12;31(1):10635. [doi: 10.1609/aaai.v31i1.10635]
39. Liehner G, Hick A, Biermann H, Brauner P, Ziefle M. Perceptions, attitudes and trust toward artificial intelligence — an assessment of the public opinion. In: Ahram T, Kalra J, Karwowski W, editors. Artificial Intelligence and Social Computing. Nice, France: AHFE Publications; 2023.
40. Chuan CH, Tsai WH, Cho SY. Framing artificial intelligence in American newspapers. In: Proceedings of the 2019 AAAI/ACM Conference on AI, Ethics, and Society. 2019 Presented at: AIES '19; January 27-28, 2019; Honolulu, HI p. 339-344 URL: https://dl.acm.org/doi/10.1145/3306618.3314285 [doi: 10.1145/3306618.3314285]
41. Carter N, Bryant-Lukosius D, DiCenso A, Blythe J, Neville AJ. The use of triangulation in qualitative research. Oncol Nurs Forum 2014 Sep;41(5):545-547. [doi: 10.1188/14.ONF.545-547] [Medline: 25158659]
42. Emerson RM, Fretz RI, Shaw LL. Writing Ethnographic Fieldnotes. 2nd edition. Chicago, IL: University of Chicago Press; 2011.
43. MAXQDA 2022. VERBI Software. URL: https://www.maxqda.com/about [accessed 2024-04-29]
44. Selting M, Auer P, Barth-Weingarten D, Bergmann JR, Bergmann P, Birkner K, et al. Gesprächsanalytisches transkriptionssystem 2 (GAT 2). Gesprachsforschung 2009:1617-1837 [FREE Full text]
45. Strauss A, Corbin JM. Basics of Qualitative Research: Grounded Theory Procedures and Techniques. Thousand Oaks, CA: Sage Publications; 1990:978.
46. Birks M, Chapman Y, Francis K. Memoing in qualitative research: probing data and processes. J Res Nurs 2008 Jan 01;13(1):68-75. [doi: 10.1177/1744987107081254]
47. Blandford A, Gibbs J, Newhouse N, Perski O, Singh A, Murray E. Seven lessons for interdisciplinary research on interactive digital health interventions. Digit Health 2018;4:2055207618770325 [FREE Full text] [doi: 10.1177/2055207618770325] [Medline: 29942629]
48. Poncette AS, Wunderlich MM, Spies C, Heeren P, Vorderwülbecke G, Salgado E, et al. Patient monitoring alarms in an intensive care unit: observational study with do-it-yourself instructions. J Med Internet Res 2021 May 28;23(5):e26494 [FREE Full text] [doi: 10.2196/26494] [Medline: 34047701]
49. Albanowski K, Burdick KJ, Bonafide CP, Kleinpell R, Schlesinger JJ. Ten years later, alarm fatigue is still a safety concern. AACN Adv Crit Care 2023 Sep 15;34(3):189-197. [doi: 10.4037/aacnacc2023662] [Medline: 37644627]
50. Managing alarms in acute care across the life span: electrocardiography and pulse oximetry. American Association of Critical Care Nurses. URL: https://www.aacn.org/clinical-resources/practice-alerts/managing-alarms-in-acute-care-across-the-life-span [accessed 2024-01-14]
51. Sendelbach S, Funk M. Alarm fatigue: a patient safety concern. AACN Adv Crit Care 2013;24(4):378-387. [doi: 10.1097/NCI.0b013e3182a903f9] [Medline: 24153215]
52. Oliveira AE, Machado AB, Santos ED, Almeida É. Alarm fatigue and the implications for patient safety. Rev Bras Enferm 2018;71(6):3035-3040 [FREE Full text] [doi: 10.1590/0034-7167-2017-0481] [Medline: 30517409]

XSL•FO
RenderX

53. Sinno ZC, Shay D, Kruppa J, Klopfenstein SA, Giesa N, Flint AR, et al. The influence of patient characteristics on the alarm rate in intensive care units: a retrospective cohort study. Sci Rep 2022 Dec 16;12(1):21801 [FREE Full text] [doi: 10.1038/s41598-022-26261-4] [Medline: 36526892]
54. Winters BD, Cvach MM, Bonafide CP, Hu X, Konkani A, O'Connor MF, Society for Critical Care Medicine Alarm and Alert Fatigue Task Force. Technological distractions (part 2): a summary of approaches to manage clinical alarms with intent to reduce alarm fatigue. Crit Care Med 2018 Jan;46(1):130-137. [doi: 10.1097/CCM.0000000000002803] [Medline: 29112077]
55. Turmell JW, Coke L, Catinella R, Hosford T, Majeski A. Alarm fatigue: use of an evidence-based alarm management strategy. J Nurs Care Qual 2017;32(1):47-54. [doi: 10.1097/NCQ.0000000000000223] [Medline: 27500694]

Abbreviations

AI: artificial intelligence
ICU: intensive care unit
INALO: Intelligent Alarm Optimizer for the Intensive Care Unit
PDMS: patient data management system
SOP: standard operating procedure

Edited by C Jacob; submitted 17.12.23; peer-reviewed by J Schlesinger, C Lehet; comments to author 10.01.24; revised version received 24.01.24; accepted 08.04.24; published 18.06.24.

Please cite as:
Mosch L, Sümer M, Flint AR, Feufel M, Balzer F, Mörike F, Poncette AS
Alarm Management in Intensive Care: Qualitative Triangulation Study
JMIR Hum Factors 2024;11:e55571
URL: https://humanfactors.jmir.org/2024/1/e55571
doi:10.2196/55571
PMID:38888941

©Lina Mosch, Meltem Sümer, Anne Rike Flint, Markus Feufel, Felix Balzer, Frauke Mörike, Akira-Sebastian Poncette. Originally published in JMIR Human Factors (https://humanfactors.jmir.org), 18.06.2024. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on https://humanfactors.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Usability of an App for Medical History Taking in General Practice From the Patients' Perspective: Cross-Sectional Study

Klara Albrink[1]; Dominik Schröder[1], MSc; Carla Joos[1]; Frank Müller[1], MSc, Dr med; Eva Maria Noack[1], PhD

Department of General Practice, University Medical Center Göttingen, Göttingen, Germany

Corresponding Author:
Eva Maria Noack, PhD
Department of General Practice
University Medical Center Göttingen
Humboldtallee 38
Göttingen, 37073
Germany
Phone: 49 551 39 65745
Email: evamaria.noack@med.uni-goettingen.de

Abstract

Background: A future shortage of physicians, especially in general practice, will result in an increasing workload for health care providers as a whole. Therefore, it is important to optimize patient-encounter processes to increase time efficiency related to visits. Utilizing digital tools to record patients' medical histories prior to a consultation offers great potential to achieve this goal. The collected information can be stored into the practice's electronic medical record, allowing for the general practitioner to review structured information of the patients' complaints and related medical history beforehand, thereby saving time during the encounter. However, the low usability of new digital developments in this setting often hinders implementation.

Objective: The aim of this study was to evaluate the usability of an app designed for medical history taking in general practice to capture the patients' perspective.

Methods: Between November 2021 and January 2022, we recruited 406 patients with acute complaints in one out-of-hour urgent care and seven general practice clinics. These study participants used the app during their waiting time and subsequently assessed its usability by completing the System Usability Scale (SUS), a robust and well-established 10-question survey measuring the perceived usability of products and technologies. Additionally, we collected general participant information, including age, sex, media usage, health literacy, and native language. Descriptive and inferential statistics were applied to identify patient characteristics associated with low or high SUS scores.

Results: We analyzed data from 397 patients (56.7% female, 43.3% male). The mean total SUS score was 77.8 points; 54.4% (216/397) of participants had SUS scores of 80 points or higher, indicating high usability of the app. In a multiple linear regression predicting SUS score, male sex and higher age (65 years or older) were significantly negatively associated with the SUS score. Conversely, a higher health literacy score and German as the native language were significantly positively associated with the SUS score.

Conclusions: Usability testing based on the SUS anticipates successful implementation of the app. However, not all patients will easily adapt to utilizing the app, as exemplified by the participants of older age in this study who reported lower perceived usability. Further research should examine these groups of people, identify the exact problems in operating such an app, and provide targeted solutions.

Trial Registration: German Clinical Trials Register World Health Organization Trial Registration Data Set DRKS00026659; https://trialsearch.who.int/Trial2.aspx?TrialID=DRKS00026659

(JMIR Hum Factors 2024;11:e47755) doi:10.2196/47755

KEYWORDS

digitization; application software; usability; mHealth; history of present illness; medical history taking

XSL•FO
RenderX

Introduction

As in many countries, demographic change is becoming evident in the German health care system, resulting in more complex, multimorbid patients [1] and a shortage of physicians [2]. Moreover, the proportion of older people in the population is rising steadily [3] and people tend to use medical services at a higher rate as they increase in age [4]. In Germany, one group that is particularly affected by this development are general practitioners (GPs) who are the first point of contact for patients requiring medical care and serve as the "gatekeepers" in the German health care system [5]. Approximately 80% of the German population aged 18 years and older are treated by a GP at least once a year [6]. A considerable number of GPs will retire in the upcoming years, resulting in 11,000 GP vacancies expected by 2035. These vacancies will disproportionately impact structurally weak and rural areas [7,8]. Without a sufficient workforce to replace the retired GPs and meet the greater demand for physicians, remaining GP workloads are expected to increase significantly within the next decade [9]. These developments challenge the German health care system at various levels and require attention to address the following key issues: future financing, improving allocation of resources, ensuring access to care, increasing efficiency and effectiveness of health care provision, and strengthened collaboration between providers [10].

To streamline patient care in the upcoming years, it is of importance to optimize patient-encounter processes to increase time efficiency related to visits. In this respect, digital tools offer great potential to support GPs in patient management, documentation workload, and the collection of medical history before consultation.

Digital tools designed to collect patients' medical history can ensure that information is always collected and documented thoroughly in a structured manner and with consistent quality. As many conditions can be diagnosed via a thorough medical history [11,12], these tools can be helpful in maintaining quality of care when time constraints may lead to an otherwise superficial medical history.

As part of our project titled "Digitally assisted information acquisition before medical consultation" (DASI), we developed an app for medical history taking in general practice settings. The app is used by the patient prior to the medical consultation, which could be either in the waiting area or at home. The collected information can be stored in the practice's electronic medical record and eventually be transferred to the individual electronic patient file, which statutory health insurers in Germany have been entitled to use since 2021 [13]. In the electronic patient file, patient data such as medical reports, X-rays, immunization records, and other medical data can be stored and shared among medical providers involved in the care of a particular patient by using the telematics infrastructure [14].

One advantage of the app is that the GP can review structured information of the patients' complaints and related medical history before the encounter. This is particularly helpful for patients that are unknown to the provider, those with many complaints, or those who have a comprehensive medical history. These situations are especially prevalent in out-of-hour urgent care practices. Furthermore, the tool might help patients to reflect on their conditions and enable them to better address their needs when seeing the provider. In this way, we expect that the limited consultation time can be used more efficiently.

Despite Germany's progress in digitalization within the health sector, concerns remain about the limited usability of new digital tools, hindering their full implementation. More than half of German practices see low usability as a strong obstacle to digitalization [15]. The evaluation of a digital tool's perceived usability is of special interest as it is a key determinant of performance for end users. Therefore, the aim of this study was to assess the usability of the app from the patients' perspective and to identify features in need of improvement. The broader aim is to ensure that the app is suitable for implementation in everyday practice, considering that GPs treat a broad range of patients of all ages and various educational and cultural backgrounds [16].

Methods

Study Design and Recruitment

This was a cross-sectional study conducted in Germany in one out-of-hour urgent care practice and seven GP practices to assess the usability of an app designed for medical history taking in general practice settings.

Software and Hardware

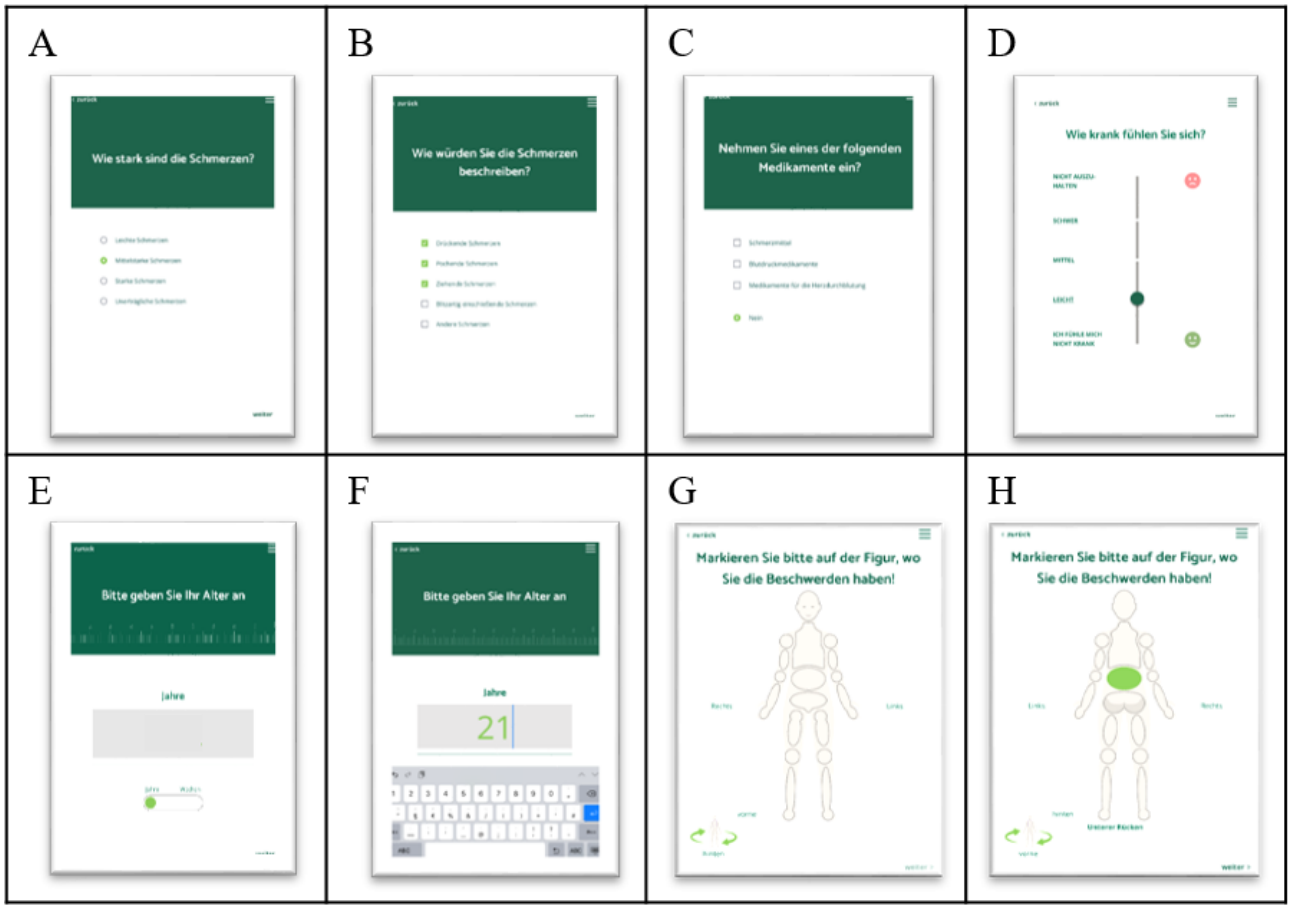
The app was developed to take a medical history based on general medical complaints directly from the patients. While there are no international standards for the composition of a standardized patient history, this app was developed based on guidelines and health literature by medical experts from aidminutes GmbH (Hamburg/Buchholz in der Nordheide, Germany). For this study, the content and query structure were further refined for primary care (general practice and out-of-hour practices) by aidminutes GmbH in collaboration with experienced researchers from the Department of General Practice at the University Medical Center Göttingen, Germany. The app was designed to be used by patients in the waiting room before they see the doctor. Patients select one or several complaints and are then guided through a symptom-related questionnaire. In the sense of a branching logic, the app is adaptive to patient responses, which trigger further specific questions about the selected key complaints (eg, how and when a symptom started). Patients are also asked about preexisting conditions, previous treatments and surgeries, current medication, living habits, and chronic conditions in the family history. Information such as biological sex, height, weight, age, as well as the subjectively perceived severity of the complaints are inquired from all patients. More details can be found in the published study protocol [17].

The app was designed to be intuitive for the user such that no prior knowledge or any kind of instruction for its use is necessary. The user interface was designed to be simple to follow and only one question is asked per screen. As the app is operated in the waiting area, sound and video output of an earlier version [18] was omitted due to data protection. The questions

XSL•FO
RenderX

are phrased in plain language; medical terminology is avoided or otherwise explained. The questions are substantially comprised by single-choice or multiple-choice questions that can be answered by tapping but also include several data fields (for age, height, and weight) and slider-type questions (Figure 1). The color scheme was designed to ensure reading accessibility for patients who may be color blind. A zoom function can be used for users who may experience visual impairment.

Figure 1. Screenshots of the app for medical history taking in general practice showing different types of questions: (A) single-choice question; (B) multiple-choice question; (C) hybrid question (ie, patients can either select several options or negate all of them); (D) slider for questions including a ranking between items (depicted here as "How sick do you feel?"); (E, F) data entry field (here: "Please enter your age"); and (G, H) selection of a body region on a figure (depicted in figure: "Please mark on the figure where you are suffering from the problems").



As this is a web-based app, it relies on a permanent internet connection. For this study, the app ran on an iPad Mini 4 (Apple Inc, Cupertino, CA, USA) held in an upright position. Tablets were equipped with haptically and visually inconspicuous cases (dark grey polyurethane leather outside and microfiber inside).

Setting

In Germany, GP practices aim at providing preventive, acute, and rehabilitative health care with long-lasting patient-doctor relationships. Out-of-hours urgent care practices provide urgent medical care for acute but not life-threatening cases when other practices are closed. Urgent care practices are often staffed with doctors of various specialties and an established relationship of care between the patients and doctors is not common. These aspects can lead to challenges in efficiently obtaining an accurate medical history and identifying serious health problems. Although the app was designed for general practice, it is also suitable to be used in out-of-hour urgent care practices.

Data Collection

The recruitment of patients was carried out by three study nurses and took place from November 22, 2021, to January 12, 2022. Patients were approached by study nurses in the waiting room of the respective practices before seeing their GP.

Patients meeting the following criteria were eligible to participate in the study: (1) seeking care in a participating practice because of acute somatic and/or psychological complaints, (2) at least 18 years old, and (3) consenting to participate in the study. Patients meeting the following exclusion criteria could not participate in the study: (1) younger than 18 years old (legal minor), (2) patients in an apparent emergency, (3) patients who required immediate medical treatment, and (4) patients who were unable to provide consent.

After the study nurses obtained written informed consent, a tablet on which the app was run on was handed over to the study participants. Participants used the app to report their medical history without an introduction on how to navigate the app. Once finished, they were asked to answer questions on personally perceived usability, media usage, and further

XSL•FO
RenderX

sociodemographic data, which were digitally attached to the medical history–taking document. The study nurse in charge was present to observe any problems study participants may have had with using the app and was available to answer questions about the app's content and usability if specifically requested. Data were collected in an anonymized format without any personal information (eg, name or address) linking the results to each study participant. More detailed information on the data collection can be found in the study protocol [17].

Ethical Considerations

The Medical Ethics Committee of the University Medical Center Göttingen approved the study (approval number 26/3/21). A written informed consent form was collected from all patients before their inclusion in the study. Participating in the study was voluntary for patients. Patients could withdraw from participation without giving a reason at any time before they had completed the survey. Subsequently, their data could no longer be deleted because it could not be traced back to the individual.

Measures

The main outcome "usability" was measured using the System Usability Scale (SUS) [19], a commonly used instrument for this purpose [20]. The SUS was developed based on Standard ISO 9241-11 [21], in which usability is measured by the three main attributes of "effectiveness," "efficiency," and "satisfaction" [22,23]. Compared to other instruments, the SUS offers several advantages: (1) it can be analyzed quickly, (2) it is relatively easy to understand by academics from other disciplines [24], (3) it contains only 10 statements for easy completion, and (4) it can be used to evaluate almost any type of user interface [25]. We used the translated and validated German version of the SUS [26] and modified the statements to suit our purpose (see Multimedia Appendix 1).

The SUS consists of 10 statements (Table 1), where statements 1, 3, 5, 7, and 9 are positively connoted and statements 2, 4, 6, 8, and 10 are negatively connoted [19]. The scores for these statements are therefore inverted when calculating the sum. The raters decide on the extent to which they agree or disagree to these statements on a 5-point Likert scale ranging from 0 (strongly disagree) to 4 (strongly agree). The final sum score is multiplied by 2.5, resulting in a score range of 0-100 with higher scores indicating better usability [19].

Lewis and Sauro [27] developed a curved grading scale for SUS scores by comparing more than 200 industrial usability studies and using the percentile ranges, resulting in grades "C" (scores of 62.7-72.5), "B" (scores of 72.6-78.8), and "A" (scores of 78.9-100). As a SUS score of 80 proves an above-average user experience, it has become a common industrial goal. This threshold was therefore used for interpreting our results.

Table 1. Items of the System Usability Scale (SUS) [19].

Items	English version of the statement
SUS 1	I think that I would like to use this system frequently
SUS 2[a]	I found the system unnecessarily complex
SUS 3	I thought the system was easy to use
SUS 4[a]	I think that I would need the support of a technical person to be able to use this system
SUS 5	I found the various functions in this system were well integrated
SUS 6[a]	I thought that there was too much inconsistency in this system
SUS 7	I would imagine that most people would learn to use this system very quickly
SUS 8[a]	I found the system very cumbersome to use
SUS 9	I felt very confident using the system
SUS 10[a]	I needed to learn a lot of things before I could get going with this system

[a]The scores of negatively connoted SUS items were inverted when calculating the sum.

Covariates

Consultations in general practice are attended by patients of different ages and educational as well as cultural backgrounds, who have a different quantities of digital interactions in everyday life. To determine whether these factors have an influence on the personally perceived usability, we surveyed age, sex, media usage, health literacy, and native language. Information about age and sex were part of the app-taken medical history. In addition to the SUS, we asked patients about which digital media tools were available to them in everyday life (possible answers: cell phone/smartphone, computer/laptop/notebook, tablet, television, none, and others; multiple answers were possible) and how many hours a day they used digital media (possible answers: 0-≤1, 1-≤2, 2-≤3, 3-≤4, or 4 or more hours). We asked three questions concerning health literacy as a proxy for education attainment, given that educational achievement is the central determinant of health literacy [28]. Questions covering the three aspects of finding/accessing, evaluating/appraising, and understanding health-related information and content were derived from the European health literacy survey [29,30] adapted for the German language (HLS-GER 2 [31]). The HLS-GER 2 uses a predefined 4-point Likert scale.

XSL•FO
RenderX

Statistical Analysis

Data from the app were saved into a database and subsequently exported to a tab separated format for further analyses. Participants with two or more missing values of the SUS questionnaire were excluded from statistical analysis. In the case of one missing SUS response, we substituted the missing value with a neutral score of 2, as this method has been used with the SUS in previous research [32].

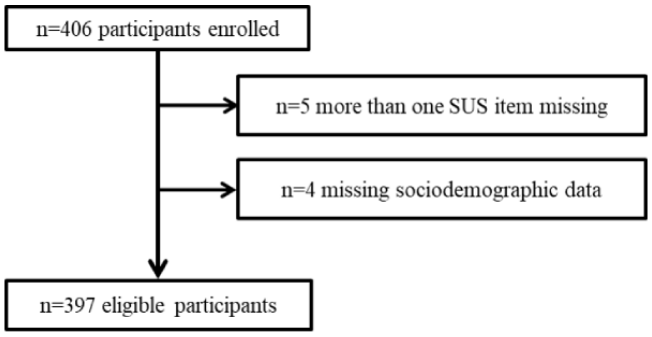
Sociodemographic data are presented as number and percentage of patients for each categorical data point. Mean and SD were utilized for interval or ratio-scaled data, which has become a common industrial goal. Sociodemographic data were compared between participants with SUS scores <80 and ≥80 using the Fisher exact test for 2×2 tables or the Fisher-Freeman-Halton test for categorical variables and the Wilcoxon rank-sum test for continuous variables. A multiple linear regression was conducted using sex, age, native German language, health literacy score, media usage duration per day, sickness level of the participants, and number of stated complaints in the app as independent variables and the SUS score as the dependent variable. Additionally, the individual SUS items were compared according to sex, age (<65 years vs ≥65 years), German native language, and tablet usage with the Wilcoxon rank-sum test. Data are visually presented as boxplots and radar charts. All analyses were carried out using R (4.1.3 under a GNU license) with the packages fmsb [33], psych [34], tidyr [35], dplyr [36], and ggplot2 [37].

Results

Patient Characteristics

We aimed to include approximately 400 patients for this study. This target was set to be able to form subgroups and to ensure that all types of patient complaints were included in our sample, including those selected on a limited basis. In total, individual data from 397 participants were included, with 5 participants having one missing SUS item. Figure 2 shows the flowchart of included patients and Table 2 shows the patients' characteristics.

Figure 2. Flowchart of patient inclusion in the cross-sectional study capturing patients' perceived usability of the app. SUS: System Usability Scale.



XSL•FO
RenderX

Table 2. Characteristics of the participants of the cross-sectional study capturing patients' perceived usability of the app (N=397).

Characteristics	Value
Sex, n (%)	
Female	225 (56.7)
Male	172 (43.3)
Age (years), median (IQR)	35.0 (25.0)
Age group (years), n (%)	
<30	152 (38.3)
30-65	223 (55.4)
65+	22 (6.3)
Native language German, n (%)	328 (82.6)
Devices used regularly[a], n (%)	
Smartphone	389 (98.0)
Tablet	210 (52.9)
Computer/notebook	310 (78.1)
Television	296 (74.6)
Media usage duration per day (hours), n (%)	
<2	89 (22.4)
2-4	174 (43.8)
>4	134 (33.8)
Self-assessed health literacy, median (IQR)[b]	
Understanding doctor	2.0 (0.0)
Search and understand health information	2.0 (1.0)
Evaluate health information	1.0 (1.0)
"How sick do you feel?"[c], n (%)	
I don't feel sick	32 (8.1)
Just a little	70 (17.6)
Fairly	226 (56.9)
Very	61 (15.4)
Unbearably	6 (1.5)

[a]Multiple selection possible.

[b]Measured on a 4-point (0-3) Likert-scale (higher scores indicate higher health literacy levels).

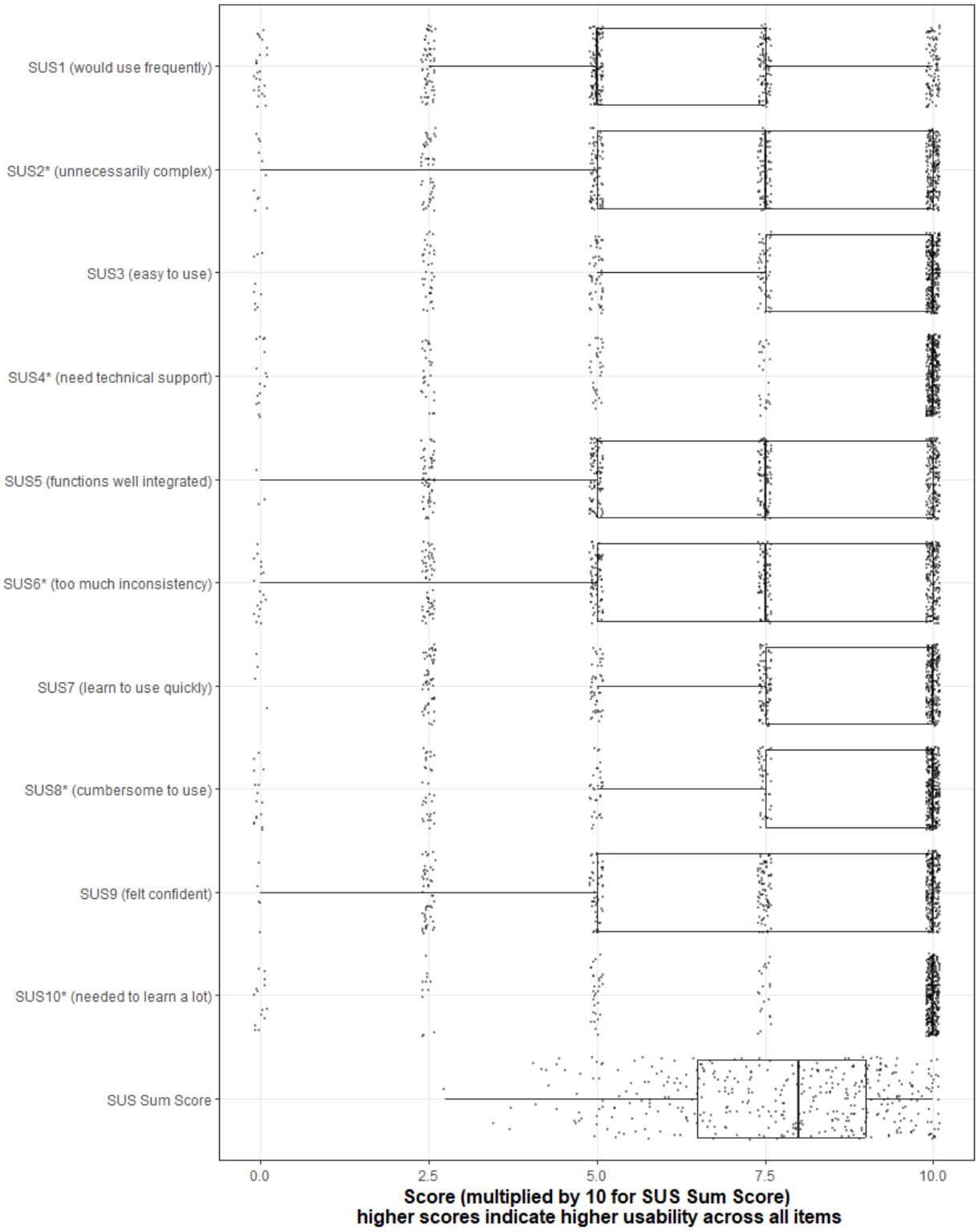
[c]Perceived severity of acute complaint.

Usability for All Participants

We found a mean total SUS score of 77.8 points, with 54.4% (216/397) of participants having SUS scores of 80 points or higher, indicating high usability of the app overall. Figure 3 shows boxplots of the individual items in which the scores were calculated for each statement. Irrespective of a positive or negative connotation, a higher score indicates a better result. The maximum score that can be achieved for each item is 10.

XSL•FO
RenderX

Figure 3. SUS items and sum scores for all study participants (N=397). SUS: System Usability Scale. *Scoring was inverted.



Usability Stratified by Sociodemographic Factors

We divided the sample into two groups with the cutoff at a SUS score of 80. Participants with a SUS score of at least 80 were significantly younger, reported higher levels of technology device usage, and higher levels of self-assessed health literacy compared to participants with a SUS score below 80 (Table 3).

XSL•FO
RenderX

Table 3. Sociodemographic variables of study participants stratified by System Usability Scale (SUS) score.

Variable	SUS score<80 (n=181)	SUS score≥80 (n=216)	*P* value
Sex, n (%)			.09[a]
Female	94 (51.9)	131 (60.6)	
Male	87 (48.1)	85 (39.4)	
Age (years), median (IQR)	38.0 (26.0)	32.5 (22.0)	.002[b]
Age group (years), n (%)			.003[c]
<30	55 (30.3)	97 (44.9)	
30-65	111 (60.2)	112 (51.4)	
65+	15 (9.4)	7 (3.7)	
Native language German, n (%)	142 (78.5)	186 (86.1)	.05[a]
Devices used regularly[d], n (%)			
Smartphone	174 (96.1)	215 (99.5)	.03[a]
Tablet	84 (46.4)	126 (58.3)	.02[a]
Computer/notebook	136 (75.1)	174 (80.6)	.22[a]
Television	125 (69.1)	171 (79.2)	.03[a]
Media usage duration per day (hours), n (%)			.08[c]
<2	48 (26.5)	41 (19.0)	
2-4	69 (38.1)	105 (48.6)	
>4	64 (35.4)	70 (32.4)	
Self-assessed health literacy, median (IQR)			
Understanding doctor	2.0 (0.0)	2.0 (1.0)	.16[b]
Search and understand health information	2.0 (0.0)	2.0 (1.0)	.01[b]
Assess confidence of health information	1.0 (1.0)	1.0 (1.0)	.19[b]
"How sick do you feel?"[e], n (%)			.35[c]
I don't feel sick	11 (6.1)	21 (9.7)	
Just a little	30 (16.6)	40 (18.5)	
Fairly	102 (56.4)	124 (57.4)	
Very	34 (18.8)	27 (12.5)	
Unbearably	3 (1.7)	3 (1.4)	

[a]Fisher exact test.
[b]Wilcoxon rank-sum test.
[c]Fisher-Freeman-Halton test.
[d]Multiple selection possible.
[e]Perceived severity of acute complaint.

A multiple linear regression predicting the SUS score was conducted, including sex, age, native German language, health literacy score, media usage duration per day, sickness level of the participants, and number of stated complaints in the app as independent variables (see Table 4). Age, sex, health literacy score, and German native language were significantly associated with SUS score. A higher age (t_{385}=3.30, *P*=.001) and male sex (t_{385}=1.98, *P*=.05) were negatively associated with SUS score, whereby a higher health literacy score (t_{385}=2.83, *P*=.01) and German as a native language (t_{385}=2.51, *P*=.01) were positively associated with SUS score.

XSL•FO
RenderX

Table 4. Multiple linear regression predicting the System Usability Scale sum score.

Variable	β (95% CI)	*P* value
Male sex (reference=female)	–3.21 (–6.40 to –.02)	.05
age (per year)	–.17 (–.27 to –.07)	.001
German not native language (reference=yes)	–5.39 (–9.61 to –1.17)	.01
Does not use tablet (reference=yes)	–1.44 (–4.66 to 1.79)	.38
Average daily media usage (reference=<2 h)		
2-4 h	3.21 (–.95 to 7.37)	.13
>4 h	.01 (–4.46 to 4.48)	.99
Health literacy score (scale 0-9)[a]	1.48 (.45 to 2.51)	.01
How sick do you feel? (score 1-5)[b]	–1.05 (–2.92 to .82)	.27
Number of stated complaints (1-11)	–.98 (–2.11 to .15)	.09

[a]Higher scores indicate a higher level of health literacy.

[b]Perceived severity of acute complaint; higher values indicate a higher level of discomfort.

Differences in Individual Items of the SUS

Stratified according to sex, age, native language, and tablet usage (see Figure 4), significant differences were detected in SUS items 2 (“unnecessarily complex”), 4 (“need technical support”), 7 (“learn to use quickly”), 8 (“cumbersome to use”), and 10 (“needed to learn a lot”).

In comparing female and male respondents, all statements were rated more positively by female participants, except for items 1 (“would use frequently”) and 4 (“need technical support”). Female participants also scored significantly higher than male participants on items 2 (“unnecessarily complex”) (mean 7.82 vs 7.11; *P*=.02), 7 (“learn to use quickly”) (mean 8.11 vs 7.53; *P*=.05), 8 (“cumbersome to use”) (mean 8.57 vs 8.08; *P*=.05), and 10 (“needed to learn a lot”) (mean 9.14 vs 8.63; *P*=.03).

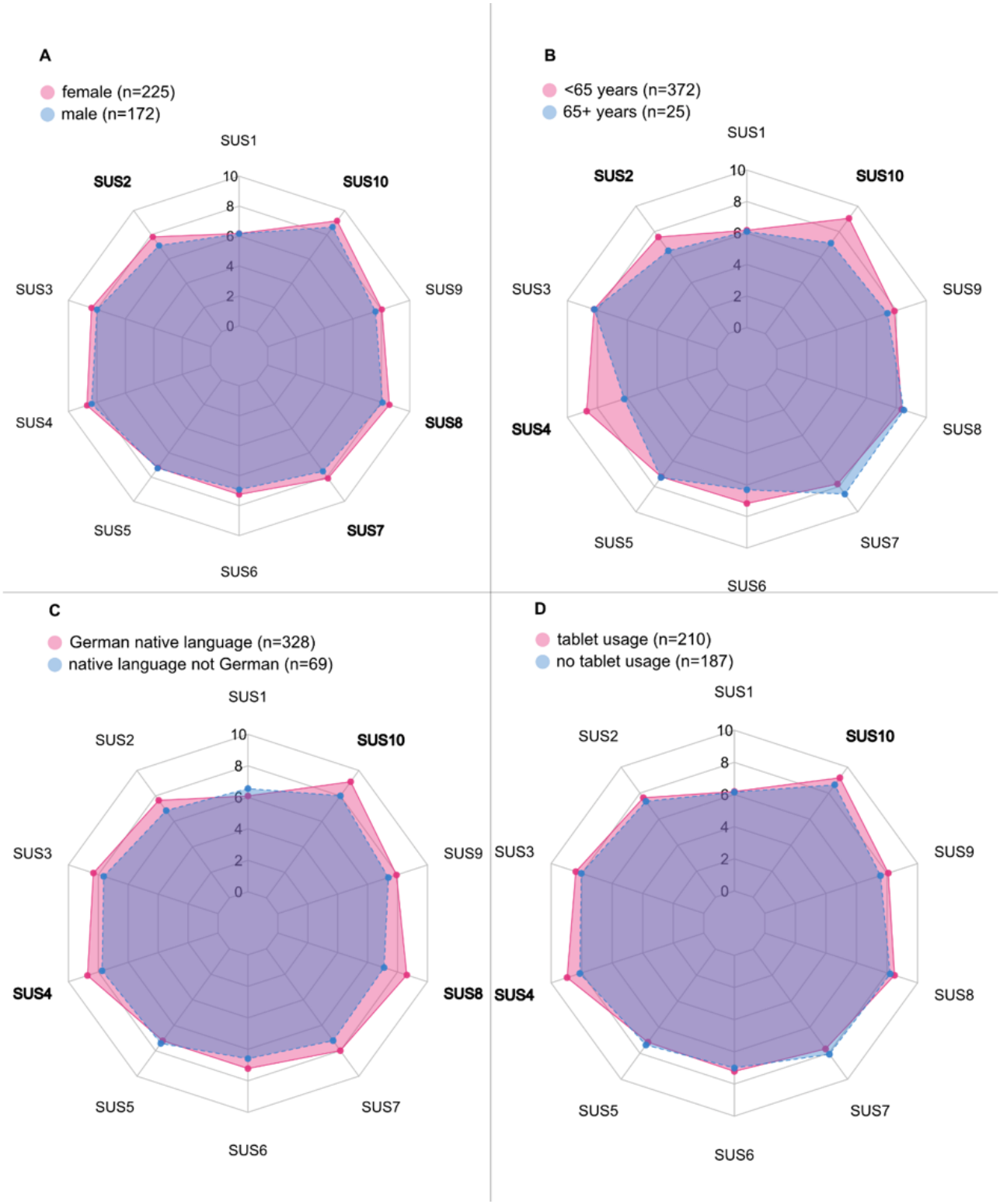
Respondents aged 65 years and older scored significantly higher on items 2 (“unnecessarily complex”) (mean 7.58 vs 6.50; *P*=.04), 4 (“need technical support”) (mean 8.72 vs 6.20; *P*<.001), and 10 (“needed to learn a lot”) (mean 9.05 vs 7.; *P*<.001) compared to their counterparts.

German native language speakers scored significantly higher on items 4 (“need technical support”) (mean 8.73 vs 7.75; *P*=.001), 8 (“cumbersome to use”) (mean 8.62 vs 7.10; *P*<.001), and 10 (“needed to learn a lot”) (mean 9.11 vs 8.01; *P*<.001) relative to nonnative speakers.

Lastly, patients who regularly use a tablet had significantly higher SUS scores on items 4 (“need technical support”) (mean 8.95 vs 8.11; *P*<.001) and 10 (“needed to learn a lot”) (mean 9.18 vs 8.64; *P*=.02) in comparison to those of participants who reported reduced levels of tablet use.

XSL•FO
RenderX

Figure 4. Mean System Usability Scale (SUS) score items (see Table 1) stratified according to sex (A), age (B), German as native language (C), and current tablet usage (D). bold: *P*<.05 (Wilcoxon rank-sum test).



Discussion

Principal Findings

In this study, we evaluated the usability of an app in taking medical histories in general practice directly from patients using the SUS [19].

The app achieved a mean SUS score of 77.9, which corresponds to a B+ grade on the curved grading scale [27] and represents a "better" product that does not necessarily need improvement [25]. Other medical devices, even those widely used at home, have lower SUS scores. Kortum and Peres [38] assessed the usability of home health care devices among students, thus representing relatively young, healthy, and well-educated participants. SUS scores for these devices ranged from 65 for

XSL•FO
RenderX

an epinephrine injector to 67 for a pregnancy test kit and 81 for a thermometer, even with previous experience using these devices.

To ensure patients can be active participants in the digital medical history–taking process, the app must be easy and intuitive to use without technical introduction or support. This importance is reflected in items 3 (“easy to use”), 4 (“need technical support“), 7 (“learn to use quickly”), and 10 (“needed to learn a lot”). Mean values between 7.8 and 8.9 for these items indicate that intuitive use has been successfully addressed in the development of our app. Item 1, assessing the frequency of app usage, scored the lowest (6.2), which can be explained by the app’s implementation solely in a medical setting and not utilized regularly in leisure time. As such, this finding is the least meaningful for our purpose.

In a pilot study by Melms et al [39], a self-completed tablet-based digital questionnaire designed for collecting medical histories in an emergency department was found to score high with respect to perceived usability. The design and content were similar to those of our app; however, their questions were only based on the SUS, which does not allow direct comparison. Other comparable instruments, although also for emergency departments, have been tested for usability in pilot studies using self-developed satisfaction surveys [40,41], a single question, and researcher or staff documentation of a patient’s need for assistance [42]. In these studies, patients were mostly satisfied with the self-administered medical history–taking tools and reported good ease of use. Taken together, these results give hope that it is possible to design a medical history app that is perceived as user-friendly.

Nonetheless, obstacles to implementing a digital tool in general practice settings can be multifaceted. Surprisingly, we found that sex was significantly associated with usability; female participants had significantly higher SUS scores than male participants. The fact that men scored higher than women for item 4 (“need technical support”) suggests that men felt more confident than women with using the app. Previous studies demonstrated that men tend to report overconfidence in their abilities, especially in fields with a male connotation [43], which computer science certainly represents [44]. Therefore, it is unclear whether men really would have needed less help or whether they overestimated themselves in their technical skills.

Our study suggests that older people are more likely to have difficulties with the handling of such an app. This aligns with a study showing that from the retirement age of 65 years, digital media use among the German population begins to decline dramatically [45] and a positive attitude toward digitalization decreases with increasing age [46]. Older age has a negative impact on the broad usability score given to a user interface [25]. To that end, this study cannot definitively conclude if the older participants of this study actually perceived the app to be of relatively low usability or if their more negative attitude toward the benefit of new technologies prompted them to give lower scores. Due to the small sample size of participants aged 65 years and older, it is not possible to assert how older people in general would cope with the handling of the app. Since GPs are consulted predominantly by older people [47], further research should focus on app testing with older patients to obtain specific feedback, including suggestions for improvement.

Having learned German as a native language was positively associated with a higher SUS score, although only patients with sufficient German language proficiency were included in the study [17]. This could be due to two different reasons: despite the app’s plain language, it is possible that some of the medical history questions or SUS items were not understood properly.

Daily media use was not associated with the SUS score, which suggests that the app is designed to also ensure that people with limited digital experience do not feel overstrained with its operation.

Limitations

Despite our efforts, this study comes with several limitations. The number of older participants (ie, aged 65 years and above) was relatively low in comparison to their constituents in GP practice settings [47]. One potential reason could be a more pronounced skepticism toward digital tools in older generations, leading to an increase in refusal for participation in the study among older patients. However, as no screening lists were maintained, this is mere speculation. A screening list should be obtained in future studies to be able to characterize individuals who declined participation. Another consideration is that people with lower levels of digital media literacy use may not have agreed to participate in the study.

Data collection was performed during the SARS-CoV-2 pandemic, which may have disproportionately impacted study participants as certain patient groups may have avoided seeing a doctor or were more likely to refuse to participate in the study to avoid unnecessary contact. This could have included especially vulnerable groups such as older people or those with multimorbid conditions.

The Likert scale of the SUS questions shown with clickable singular dots was replaced by a slider on December 8, 2021. In the dot-based representation, it was compulsory to make an entry before continuing, whereas the slider was automatically set to the neutral center and could be shifted. This may have led to incorrectly rated items. For example, this may have occurred in instances of the internet faltering or the patient having double-clicked without noticing. Since there were repeated questions about the word “Inkonsistenzen” (SUS item 2), we replaced it by the more common synonym “Unstimmigkeiten” (English translation: discrepancies). Furthermore, hardware as well as the operating system may have influenced the evaluation of the personally perceived usability of a system [48]. For this study, iPad Minis with the iOS operating system were exclusively used. Therefore, possible differences in assessment related to the operating system and hardware are not part of this study.

The SUS is able to classify the usability of a system but is unable to identify specific usability issues nor capture the usability of the system in its entirety. For a more in-depth usability evaluation, different methods could be used (eg, interviews and observations). During data collection, staff were able to observe usability problems. In their observations, multiple-choice, single-choice, and hybrid questions as well as the slider did not

XSL•FO
RenderX

appear to cause any difficulties. In contrast, problems concerning the handling of the app arose when participants were required to input free-text entries (eg, age, height, weight). Further, some study participants were unclear on how to open and close the on-screen keyboard. Some participants also did not understand that the figure on which a pain or an injury could be assigned to a body region (see Figure 1E) could be rotated by clicking on an icon at the bottom left of the screen. This means that, for example, back pain may have been falsely reported as abdominal pain. Lastly, an unstable internet connection arose during data collection, which caused the app to be unresponsive intermittently. These factors may have influenced the SUS score.

Conclusion

The app examined in this study for medical history taking passes the usability test based on the SUS and appears to function on par with other digital tools that have become well-integrated in our everyday lives. However, not all people adapted equally well to the app. For successful implementation, all end users, regardless of age, technical affinity, health literacy, or preferred language, must be able to use such a tool. Only if that is attained, providing practical digital solutions can contribute to the efficient and effective delivery of health care services. Therefore, further research should focus on the identification of causes for difficulties of using the app as well as finding appropriate solutions.

Acknowledgments

We would like to thank Julie Ngo, MD for proofreading. We sincerely thank all patients for their participation. We are very grateful to the physicians and their teams who allowed data collection in their practices despite the SARS-CoV-2 pandemic and the extraordinary burden that came along for health care workers. We thank the team at aidminutes GmbH for their commitment and support. This research is funded by the German Innovation Fund (funding number 01VSF19050) of the Federal Joint Committee (G-BA). The funders have not influenced the design of this study and did not play any role during its implementation (eg, data collection and analysis, interpretation, and publication of the results).

Data Availability

The data sets used and analyzed during the study are not publicly available due to the decision of the research ethics board but can be obtained from the authors upon reasonable request within a data sharing agreement.

Authors' Contributions

KA wrote the original draft. KA and EMN were mainly responsible for writing the manuscript. KA, EMN, and DS conceived the study design and analyzed the data. FM and CJ revised the manuscript. All authors have read and approved the final version of the manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Original and German versions of the System Usability Scale (SUS) and the items customized for this study.
[DOCX File , 15 KB - humanfactors_v11i1e47755_app1.docx]

References

1. GBE kompakt 3(2) Demografische Alterung und Folgen für das Gesundheitswesen. Robert Koch Institut. 2012. URL: https://www.rki.de/DE/Content/Gesundheitsmonitoring/Gesundheitsberichterstattung/GBEDownloadsK/2012_2_Demografischer_Wandel_Alterung.pdf?__blob=publicationFile [accessed 2024-01-02]
2. Gesundheit in Deutschland. Gesundheitsberichterstattung des Bundes. 09 Welche Auswirkungen hat der demografische Wandel auf Gesundheit und Gesundheitsversorgung? Robert Koch-Institut. 2015. URL: https://www.rki.de/DE/Content/Gesundheitsmonitoring/Gesundheitsberichterstattung/GBEDownloadsGiD/2015/09_gesundheit_in_deutschland.pdf?__blob=publicationFile [accessed 2024-01-02]
3. Die Bevölkerungsgruppe der älteren Menschen ab 65 Jahren 2021. Statistisches Bundesamt. URL: https://www.destatis.de/DE/Themen/Querschnitt/Demografischer-Wandel/Aeltere-Menschen/bevoelkerung-ab-65-j.html [accessed 2022-05-19]
4. Kürschner N, Weidmann C, Müters S. Who enrolls in a general practitioner model? The behavioral model of health services use and general practitioner-centered care in Germany. Bundesgesundheitsblatt Gesundheitsforschung Gesundheitsschutz 2011 Feb;54(2):221-227. [doi: 10.1007/s00103-010-1205-x] [Medline: 21290277]
5. Schnitzer S, Balke K, Walter A, Litschel A, Kuhlmey A. Do gatekeeping programs increase equality of health care in Germany? A comparison of the health care situation of participants and nonparticipants. Bundesgesundheitsblatt Gesundheitsforschung Gesundheitsschutz 2011 Aug;54(8):942-950. [doi: 10.1007/s00103-011-1317-y] [Medline: 21800242]

XSL•FO
RenderX

6. Prütz F, Rommel A, Thom J, Du Y, Sarganas G, Starker A. Inanspruchnahme ambulanter medizinischer Leistungen in Deutschland – Ergebnisse der Studie GEDA 2019/2020-EHIS. J Health Monitor 2021;6(3):49-71. [doi: 10.25646/8554]
7. GESUNDHEITSDATEN: Niedergelassene Ärztinnen und Ärzte werden immer älter. Kassenärztliche Bundesvereinigung (KBV). 2022. URL: https://gesundheitsdaten.kbv.de/cms/html/16397.php [accessed 2022-05-19]
8. Beerheide R, Haserück A. Versorgungsstrukturen: Hausärztemangel angehen. Deutsches Ärzteblatt 2021;118(38):A-1684 /B-1396.
9. Schulz M, Czihal T, Bätzing-Feigenbaum J, von Stillfried D. Zukünftige relative Beanspruchung von Vertragsärzten – Eine Projektion nach Fachgruppen für den Zeitraum 2020 bis 2035. Versorgungsatlas. 2016. URL: https://www.versorgungsatlas.de/fileadmin/ziva_docs/67/VA-16-02-Bericht-Final_V3.pdf [accessed 2022-05-19]
10. Blümel M, Spranger A, Achstetter K, Maresso A, Busse R. Germany: health system review. Health Syst Transit 2020 Dec;22(6):1-272 [FREE Full text] [Medline: 34232120]
11. Hampton JR, Harrison MJ, Mitchell JR, Prichard JS, Seymour C. Relative contributions of history-taking, physical examination, and laboratory investigation to diagnosis and management of medical outpatients. Br Med J 1975 May 31;2(5969):486-489 [FREE Full text] [doi: 10.1136/bmj.2.5969.486] [Medline: 1148666]
12. Roshan M, Rao AP. A study on relative contributions of the history, physical examination and investigations in making medical diagnosis. J Assoc Physicians India 2000 Aug;48(8):771-775. [Medline: 11273467]
13. Bundesgesetzblatt Jahrgang 2020 Teil I Nr. 46, ausgegeben zu Bonn am 19. Gesetz zum Schutz elektronischer Patientendaten in der Telematikinfrastruktur/Patientendaten-Schutz-Gesetz (PDSG). Bundesrepublik Deutschland. 2020 Oct. URL: https://www.bundesgesundheitsministerium.de/fileadmin/Dateien/3_Downloads/Gesetze_und_Verordnungen/GuV/P/PDSG_bgbl.pdf [accessed 2024-01-02]
14. Patientendaten-Schutz-Gesetz. Bundesministerium für Gesundheit. URL: https://www.bundesgesundheitsministerium.de/patientendaten-schutz-gesetz.html [accessed 2024-01-02]
15. Albrecht M, Otten M, Sander M. Praxisbarometer Digitalisierung 2021. KBV. URL: https://www.kbv.de/media/sp/IGES_PraxisBaroDigit2021_Kurzbericht.pdf [accessed 2022-05-19]
16. Zapata BC, Fernández-Alemán JL, Idri A, Toval A. Empirical studies on usability of mHealth apps: a systematic literature review. J Med Syst 2015 Mar;39(2):1. [doi: 10.1007/s10916-014-0182-2] [Medline: 25600193]
17. Albrink K, Joos C, Schröder D, Müller F, Hummers E, Noack EM. Obtaining patients' medical history using a digital device prior to consultation in primary care: study protocol for a usability and validity study. BMC Med Inform Decis Mak 2022 Jul 19;22(1):189 [FREE Full text] [doi: 10.1186/s12911-022-01928-0] [Medline: 35854290]
18. Müller F, Chandra S, Furaijat G, Kruse S, Waligorski A, Simmenroth A, et al. A Digital Communication Assistance Tool (DCAT) to obtain medical history from foreign-language patients: development and pilot testing in a primary health care center for refugees. Int J Environ Res Public Health 2020 Mar 20;17(4):1368 [FREE Full text] [doi: 10.3390/ijerph17041368] [Medline: 32093286]
19. Brooke J. SUS: A 'quick and dirty' usability scale. In: Jordan PW, Thomas B, McClelland IL, Weerdmeester B, editors. Usability Evaluation In Industry. London: CRC Press; 1995:189-194.
20. Lewis JR. The System Usability Scale: past, present, and future. Int J Hum Comput Interact 2018 Mar 30;34(7):577-590. [doi: 10.1080/10447318.2018.1455307]
21. Ergonomics of human-system interaction – Part 11: Usability: Definitions and concepts. International Organization for Standardization (ISO). 2018. URL: https://www.iso.org/obp/ui/#iso:std:iso:9241:-11:ed-2:v1:en [accessed 2022-05-19]
22. Abran A, Khelifi A, Suryn W, Seffah A. Usability meanings and interpretations in ISO standards. Softw Qual J 2003;11(4):325-338. [doi: 10.1023/a:1025869312943]
23. Kortum P, Hebl M, Oswald FL. Applying usability measures to assess textbooks. 2014 Presented at: 58th Human Factors and Ergonomics Society Annual Meeting; October 27-31, 2014; Chicago, IL. [doi: 10.1177/1541931214581281]
24. Banger A, Kortum P, Miller J. Determining what individual SUS Scores mean: adding an adjective rating scale. J Usability Stud 2009;4:114-123 [FREE Full text]
25. Bangor A, Kortum PT, Miller JT. An empirical evaluation of the System Usability Scale. Int J Hum Comput Interact 2008 Jul 30;24(6):574-594. [doi: 10.1080/10447310802205776]
26. Gao M, Kortum P, Oswald FL. Multi-language toolkit for the System Usability Scale. Int J Hum Comput Interact 2020 Aug 19;36(20):1883-1901. [doi: 10.1080/10447318.2020.1801173]
27. Lewis JR, Sauro J. Item benchmarks for the System Usability Scale. J Usability Stud 2018;13(3):158-167 [FREE Full text]
28. Sørensen K, Pelikan JM, Röthlin F, Ganahl K, Slonska Z, Doyle G, HLS-EU Consortium. Health literacy in Europe: comparative results of the European health literacy survey (HLS-EU). Eur J Public Health 2015 Dec;25(6):1053-1058 [FREE Full text] [doi: 10.1093/eurpub/ckv043] [Medline: 25843827]
29. Sørensen K, Van den Broucke S, Fullam J, Doyle G, Pelikan J, Slonska Z, (HLS-EU) Consortium Health Literacy Project European. Health literacy and public health: a systematic review and integration of definitions and models. BMC Public Health 2012 Jan 25;12:80 [FREE Full text] [doi: 10.1186/1471-2458-12-80] [Medline: 22276600]
30. HLS-EU CONSORTIUM. Comparative Report of Health Literacy in Eight EU Member States (Second Revised and Extended Version): The European Health Literacy Survey HLS-EU. 2014 Jul 22. URL: https://www.hsph.harvard.edu/wp-content/uploads/sites/135/2015/09/neu_rev_hls-eu_report_2015_05_13_lit.pdf [accessed 2022-05-19]

XSL•FO
RenderX

31. Schaeffer D, Gille S, Berens E, Griese L, Klinger J, Vogt D, et al. Digital health literacy of the population in Germany: results of the HLS-GER 2. Gesundheitswesen 2023 Apr 14;85(4):323-331. [doi: 10.1055/a-1670-7636] [Medline: 34905785]
32. Downey RG, King C. Missing data in Likert ratings: a comparison of replacement methods. J Gen Psychol 1998 Apr;125(2):175-191. [doi: 10.1080/00221309809595542] [Medline: 9935342]
33. Package 'fsmb'. CRAN R project. 2023. URL: https://cran.r-project.org/web/packages/fmsb/fmsb.pdf [accessed 2023-12-29]
34. Revelle WR. psych: procedures for personality and psychological research. Northwestern University Northwestern Scholars. 2022. URL: https://www.scholars.northwestern.edu/en/publications/psych-procedures-for-personality-and-psychological-research [accessed 2022-05-19]
35. tidyr: Tidy Messy Data. URL: https://tidyr.tidyverse.org/ [accessed 2022-05-19]
36. dplyr: A Grammar of Data Manipulation. URL: https://dplyr.tidyverse.org/ [accessed 2023-12-29]
37. Wickham H. ggplot2: Elegant graphics for data analysis, second edition. Cham: Springer; 2016.
38. Kortum P, Peres SC. Evaluation of home health care devices: remote usability assessment. JMIR Hum Factors 2015 Jun 05;2(1):e10 [FREE Full text] [doi: 10.2196/humanfactors.4570] [Medline: 27025664]
39. Melms L, Schaefer JR, Jerrentrup A, Mueller T. A pilot study of patient satisfaction with a self-completed tablet-based digital questionnaire for collecting the patient's medical history in an emergency department. BMC Health Serv Res 2021 Jul 30;21(1):755 [FREE Full text] [doi: 10.1186/s12913-021-06748-y] [Medline: 34330279]
40. Arora S, Goldberg A, Menchine M. Patient impression and satisfaction of a self-administered, automated medical history-taking device in the Emergency Department. West J Emerg Med 2014 Mar;15(1):35-40 [FREE Full text] [doi: 10.5811/westjem.2013.2.11498] [Medline: 24695871]
41. Benaroia M, Elinson R, Zarnke K. Patient-directed intelligent and interactive computer medical history-gathering systems: a utility and feasibility study in the emergency department. Int J Med Inform 2007 Apr;76(4):283-288. [doi: 10.1016/j.ijmedinf.2006.01.006] [Medline: 16473548]
42. Herrick DB, Nakhasi A, Nelson B, Rice S, Abbott PA, Saber Tehrani AS, et al. Usability characteristics of self-administered computer-assisted interviewing in the emergency department: factors affecting ease of use, efficiency, and entry error. Appl Clin Inform 2013;4(2):276-292 [FREE Full text] [doi: 10.4338/ACI-2012-09-RA-0034] [Medline: 23874364]
43. Jakobsson N, Levin M, Kotsadam A. Gender and overconfidence: effects of context, gendered stereotypes, and peer group. Adv Appl Sociol 2013;3(2):137-141. [doi: 10.4236/aasoci.2013.32018]
44. Friedrich JD, Hachmeister CD, Nickel S, Peksen S, Roessler I. Frauen in Informatik: Welchen Einfluss haben inhaltliche Gestaltung, Flexibilisierung und Anwendungsbezug der Studiengänge auf den Frauenanteil? CHE Arbeitspapier Nr 200. Centrum für Hochshulentwicklung (CHE). 2018 Apr. URL: https://www.uni-bremen.de/fileadmin/user_upload/sites/zentrale-frauenbeauftragte/Berichte/CHE_AP_200_Frauen_in_Informatik.pdf [accessed 2022-05-19]
45. Digital skills gap. Initiative D21. 2021. URL: https://initiatived21.de/publikationen/digital-skills-gap [accessed 2022-05-19]
46. Digitale Teilhabe lässt ab 80 Jahren deutlich nach. Initiative D21. 2021. URL: https://initiatived21.de/digitale-teilhabe-laesst-ab-80-jahren-deutlich-nach/ [accessed 2022-05-19]
47. Bergmann E, Kalcklösch M, Tiemann F. Public health care utilisation. Initial results of the Telephone Health Survey 2003. Bundesgesundheitsblatt Gesundheitsforschung Gesundheitsschutz 2005 Dec;48(12):1365-1373. [doi: 10.1007/s00103-005-1167-6] [Medline: 16270187]
48. Calisir F, Cevikcan E, Camgoz AH. Industrial engineering in the big data era. Cham: Springer; 2019.

Abbreviations

DASI: digitally assisted information acquisition before medical consultation
GP: general practitioner
HLS-GER 2: European health literacy survey adapted for the German language
SUS: System Usability Scale

Edited by A Kushniruk; submitted 31.03.23; peer-reviewed by A Simmenroth, R Marshall; comments to author 27.08.23; revised version received 10.10.23; accepted 20.11.23; published 05.01.24.

Please cite as:
Albrink K, Schröder D, Joos C, Müller F, Noack EM
Usability of an App for Medical History Taking in General Practice From the Patients' Perspective: Cross-Sectional Study
JMIR Hum Factors 2024;11:e47755
URL: https://humanfactors.jmir.org/2024/1/e47755
doi:10.2196/47755
PMID:38180798

XSL•FO
RenderX

©Klara Albrink, Dominik Schröder, Carla Joos, Frank Müller, Eva Maria Noack. Originally published in JMIR Human Factors (https://humanfactors.jmir.org), 05.01.2024. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on https://humanfactors.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

A Novel Continuous Real-Time Vital Signs Viewer for Intensive Care Units: Design and Evaluation Study

Shiming Yang[1], PhD; Samuel Galvagno[1], MD, PhD; Neeraj Badjatia[2], MD; Deborah Stein[3], MD; William Teeter[4], MD; Thomas Scalea[3], MD; Stacy Shackelford[5], MD; Raymond Fang[6], MD; Catriona Miller[3], PhD; Peter Hu[1], PhD; VS viewer study group[7]

[1]Department of Anesthesiology, University of Maryland School of Medicine, Baltimore, MD, United States
[2]Department of Neurology, University of Maryland School of Medicine, Baltimore, MD, United States
[3]Department of Surgery, University of Maryland School of Medicine, Baltimore, MD, United States
[4]Emergency Medicine, University of Maryland School of Medicine, Baltimore, MD, United States
[5]United States Air Force Academy, Colorado Springs, CO, United States
[6]Department of Surgery, Johns Hopkins University School of Medicine, Baltimore, MD, United States
[7]See Acknowledgments

Corresponding Author:
Shiming Yang, PhD
Department of Anesthesiology
University of Maryland School of Medicine
11 S Paca St. LL01
Baltimore, MD, 21201
United States
Phone: 1 4103284179
Email: syang@som.umaryland.edu

Abstract

Background: Clinicians working in intensive care units (ICUs) are immersed in a cacophony of alarms and a relentless onslaught of data. Within this frenetic environment, clinicians make high-stakes decisions using many data sources and are often oversaturated with information of varying quality. Traditional bedside monitors only depict static vital signs data, and these data are not easily viewable remotely. Clinicians must rely on separate nursing charts—handwritten or electric—to review physiological patterns, including signs of potential clinical deterioration. An automated physiological data viewer has been developed to provide at-a-glance summaries and to assist with prioritizing care for multiple patients who are critically ill.

Objective: This study aims to evaluate a novel vital signs viewer system in a level 1 trauma center by subjectively assessing the viewer's utility in a high-volume ICU setting.

Methods: ICU attendings were surveyed during morning rounds. Physicians were asked to conduct rounds normally, using data reported from nurse charts and briefs from fellows to inform their clinical decisions. After the physician finished their assessment and plan for the patient, they were asked to complete a questionnaire. Following completion of the questionnaire, the viewer was presented to ICU physicians on a tablet personal computer that displayed the patient's physiologic data (ie, shock index, blood pressure, heart rate, temperature, respiratory rate, and pulse oximetry), summarized for up to 72 hours. After examining the viewer, ICU physicians completed a postview questionnaire. In both questionnaires, the physicians were asked questions regarding the patient's stability, status, and need for a higher or lower level of care. A hierarchical clustering analysis was used to group participating ICU physicians and assess their general reception of the viewer.

Results: A total of 908 anonymous surveys were collected from 28 ICU physicians from February 2015 to June 2017. Regarding physicians' perception of whether the viewer enhanced the ability to assess multiple patients in the ICU, 5% (45/908) strongly agreed, 56.6% (514/908) agreed, 35.3% (321/908) were neutral, 2.9% (26/908) disagreed, and 0.2% (2/908) strongly disagreed.

Conclusions: Morning rounds in a trauma center ICU are conducted in a busy environment with many data sources. This study demonstrates that organized physiologic data and visual assessment can improve situation awareness, assist clinicians with recognizing changes in patient status, and prioritize care.

(JMIR Hum Factors 2024;11:e46030) doi:10.2196/46030

XSL•FO
RenderX

KEYWORDS

clinical decision-making; health information technology; intensive care units; patient care prioritization; physiological monitoring; visualization; vital signs

Introduction

Clinicians working in intensive care units (ICUs) must be able to see, understand, and respond quickly to the complex and ever-changing clinical environment of the ICU. They need to be able to collect, analyze, and interpret what is happening and what it means [1]. Situational awareness is essential for ICU clinicians to provide safe and effective care to their patients. When clinicians have good situational awareness, they are better able to identify and respond to changes in their patients' condition and to coordinate care with other members of the health care team. However, clinicians are immersed in a cacophony of alarms and a relentless onslaught of data. Within this frenetic environment, clinicians make high-stakes decisions using multiple data sources and are often oversaturated with information of varying quality. While modern hospitals are equipped with bedside monitors collecting various physiological data in a real-time, continuous, and automated way, these data are not always easily accessible remotely or available to be viewed as a continuous trend [2]. The enormous amount of unprocessed data adds an additional burden on ICU clinicians who work in a dynamic environment with voluminous decision-making requirements. Traditional bedside monitors only show a single patient's instantaneous (static) vital signs (VS) data, limiting the clinician's scope to view a patient's physiological trajectory within a clinically meaningful period of time. Clinicians must rely on separate nursing charts—handwritten or electronic—to review a patient's physiological status. Moreover, auditory alarms often cause "alarm fatigue" instead of increasing situational awareness [3]. Many bedside monitors only display 1 or 2 patients' information; the ability to view an entire unit or ward allows a clinician to prioritize attention to those in most need of critical care support [4]. Improved visualization of patient information may help clinicians cope with information overload in critical care settings by improving situational awareness and supporting clinical decision-making [5]. An automated physiological data-organizing and information-summary system that presents aggregated information from multiple data sources while providing at-a-glance summaries of clinical data can assist ICU clinicians with prioritizing care for multiple patients.

Developed initially for use in aircraft transporting multiple patients who are critically ill, this VS viewer has 2 outcomes of direct and important clinical applicability. First, the VS viewer can provide clinicians with the capability to monitor individual patient trends, improving overall decision-making. Since patients in the ICU require multiple life support treatments to ensure ideal long-term outcomes, improved display of VS patterns could improve patient assessment and clinical decision-making. Second, the VS viewer system allows remote monitoring of groups of patients through a display that provides clinicians with the ability to quickly identify patients in need of rapid intervention. The objective of this work is to evaluate the use of a VS viewer in ICUs at a high-volume level 1 trauma center. We hypothesized that clinicians would subjectively report improved situational awareness and enhanced ability to make clinical decisions with the use of a VS viewer.

Methods

Data and System Design

In the ICUs of the University of Maryland Medical System, GE Marquette Solar 7000/8000 (General Electric) patient VS monitors are networked to provide a collection of real-time patient VS data streams. Each patient monitor collects real-time 240 Hz waveforms and 0.5 Hz trend data, which are transferred through the secure intranet to a dedicated BedMaster server (Excel Medical Electronics) and archived [6]. To increase the system's availability and reliability, a triple-redundant design was used, in which 3 BedMaster servers were used in parallel to collect data from all bed units [7]. Physiological data collected through this system, when they are displayed on the GE Marquette monitor, include electrocardiographic, photoplethysmographic, carbon dioxide, arterial blood pressure, and intracranial pressure (ICP), among others. Trends include heart rate (HR), respiratory rate, temperature, oxygen saturation, end-tidal carbon dioxide, and ICP, among many others. This information provides continuous VS data that relays important physiological information regarding brain perfusion, cardiac stability, overall tissue perfusion, and respiratory status.

During the design of the VS viewer for ICU, our goal was to create a novel physiological data displayer that can reduce ICU clinicians' workload, enhance clinical decision-making, and improve communication in a noisy and confined ICU environment. To achieve the goal, we considered the factors of usability and patient safety, which can be closely related in this application. For usability, current bedside monitors often suffer from insufficient time windows to display physiological trends, a lack of clear indications of patients' physiological status, and a lack of overview of multiple patients for prioritizing [4]. To enhance the clinicians' efficiency while maximizing patients' safety, we adopted the following design strategies: First, the viewer should reduce the information overload for clinicians to access patients' physiological data, current or past, individual or group [8-10]. Second, it should be compatible with the existing patient monitor system so that clinicians can reuse their existing knowledge about the monitor, which may increase the acceptance of the VS viewer [8]. Third, in the user experience design, the viewer should place the user in control [11]. It should use simple colors and graphs to convey efficient information while still providing detailed data for advanced users to access with simple operations [12]. Fourth, the viewer should have reasonable reliability for patients' safety. Redundance was introduced in the design for key components in the system, such as the data collection, database, and web server [7].

The VS viewer adopted a client-server architecture. The server handles 2 types of clients: the bedside monitors and the users. It receives and persists in real-time physiologic data that are

XSL•FO
RenderX

transmitted from the bedside monitors. A database records each bedside VS value, bed name, and timestamp. The server also responds to users' requests for viewing data within a given time frame. To continuously present the latest data to the user with low latency, the VS viewer uses the asynchronous Javascript and XML technique to pull the most recent data from the database every minute [13]. Such a method allows the VS viewer to automatically redraw all VS trajectories without refreshing the entire viewing page.

The VS viewer provides a rich interface for data monitoring, exploration, and recording. Data are depicted according to each clinical area of operations, such as the trauma resuscitation unit or emergency department, operating room, computed tomography suite, and individual critical care units. Figure 1 demonstrates the grouping of bed units. On the left panel, a list of all groups can be used as a shortcut to bed units. Selecting a specific unit, a default 24-hour view is displaced for shock index (SI=HR/systolic blood pressure), HR, systolic blood pressure, ICP, cerebral perfusion pressure, brain trauma index, and end-tidal carbon dioxide concentration. If ICP data are not collected, the space is used to plot the next available VS, optimizing the view.

When a bed is selected, a page for this bed (unit view) is displayed. Figure 2 demonstrates the structure of the information. The page is partitioned into multiple areas for navigation, viewing, and tools. Its center is assigned for presenting the selected patient's physiologic data in a time frame (up to 72 hours). VS trajectories are stacked vertically in order of predefined importance. The bottom is reserved for plotting bar segments of all VS that summarize the colored warnings without showing the value changes. This provides a summary of all available VS trends in a condensed space, which could be used to view the physiological stability of the patient over time. To provide an at-a-glance view of other rooms in this group, the left panel lists all the rooms in the current group and updates their VS trajectories in real time. The color-coded warning in the thumbnails enhances situation awareness even when the users are focusing on 1 patient.

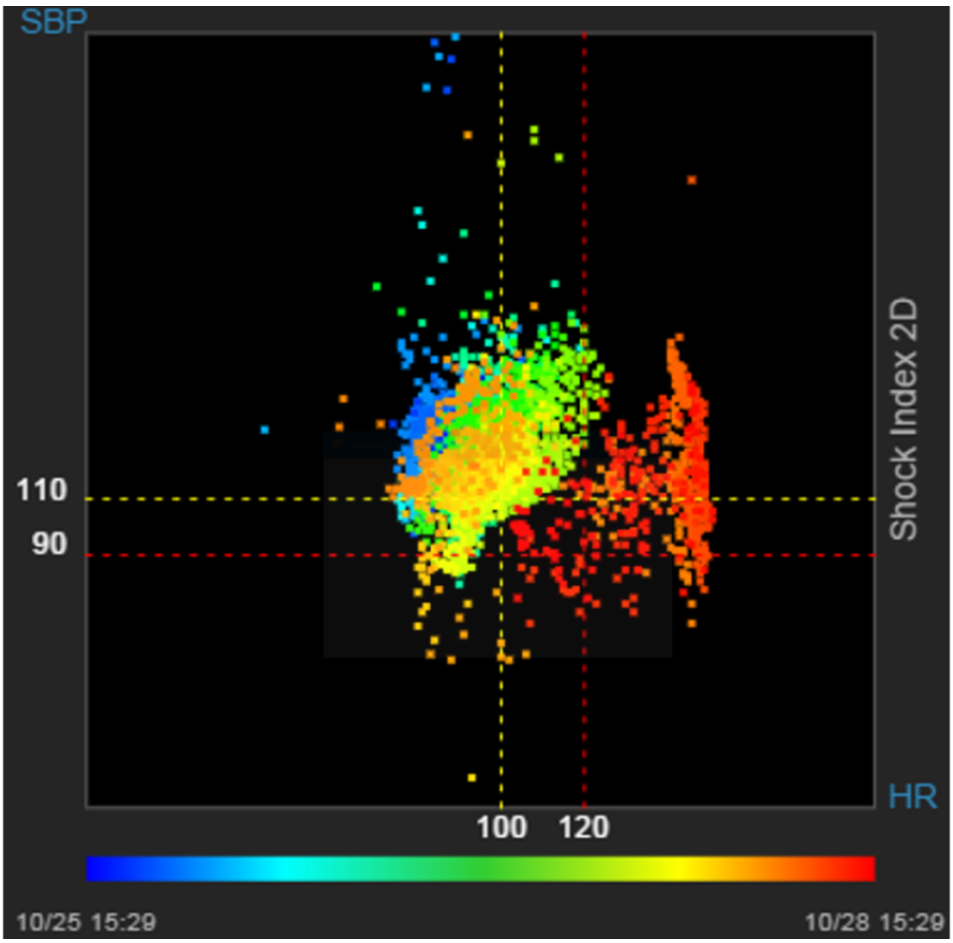
The VS viewer has additional diagnostic tools. For example, SI is a commonly used blood transfusion diagnosis tool [14]. The VS viewer adds a 2D SI diagram to show a changing trajectory (Figure 3). To present the temporal information, a heat map is plotted, ranging from blue (cold) to red (warm); blue colors represent past events, whereas red colors represent current data trends. Similarly, the brain trauma index (which is ICP or cerebral perfusion pressure) can also be visualized in the 2D plot [15].

Figure 1. The vital sign viewer in the "group" mode, with a default 24-hour display.

XSL•FO
RenderX

Figure 2. Vital sign (VS) viewer in the "unit" mode, with default 24-hour display. Labeled area 1: navigation menu to other room groups. Area 2: title information for room name, current time, and the next update time. Area 3: user portal. Area 4: list of beds in the same group with their current VS thumbnails. Area 5: the main area to display selected room VS trajectories and the summarization with color-coded patterns. Area 6: diagnostic tools for 2D scatter plots of shock index (=heart rate/systolic blood pressure) and brain trauma index (=intracranial pressure/cerebral perfusion pressure). Area 7: functional buttons for selecting various time ranges for viewing.

Figure 3. An example 2D shock index plot. The colored scatter plot shows the change in shock index (heart rate/systolic blood pressure) from past (blue) to recent (red), thereby depicting a 3-day change in worsening shock index.



XSL•FO
RenderX

Clinical Thresholds

Colored warnings are an effective means to gain a clinician's attention and may be more effective than audible alarms, especially in a noisy, busy, and confined environment [16]. In the VS viewer, VS trajectories with colors may be viewed to highlight the sections where the VS are outside of normal clinical thresholds. For example, too low or too high HR segments are displayed differently from normal HR. Clinical thresholds for VSs were developed after surveying 47 clinicians (24 medical doctors, 18 registered nurses, and 5 respiratory therapists). Among them, 36 clinicians were from the University of Maryland, Baltimore, and 11 from the University of Cincinnati. After the survey was completed, a team of clinicians met to review the results to reach a consensus on the viewer's opinion of their visual appearance. Multimedia Appendix 1 summarizes the optional threshold distributions for some important VS. Based on these threshold values, a consensus set of color-coded cutoffs was determined (Multimedia Appendix 2). These values were set as fixed parameters under consideration of a simplified and consistent user interface.

Survey Design

Clinicians who were scheduled to work in the ICU or on the trauma teams were contacted and trained on how to use the VS viewer. Once trained, ICU and team clinicians were asked to participate in the study. Clinicians were surveyed anonymously from Tuesday to Friday and were asked to conduct rounds normally, using data reported from nurse charts and briefs from fellows to inform their clinical decisions. None of those clinicians participated in the design of the VS viewer. A total of 2 questionnaires were designed to collect clinicians' opinions about a patient's condition and satisfaction with the VS viewer. Clinicians were given a preview survey upon their assessment and formulation of their plan for each patient after traditional rounds and before accessing the viewer. Immediately following the completion of the pre-view survey, the VS viewer was presented to the clinicians on a tablet, displaying the patient's past physiologic data visualized and summarized for up to 72 hours. After reviewing the viewer for up to 1 minute, clinicians completed the postview questionnaire. In both questionnaires, the clinicians answered questions regarding the patient's stability, status, and need for a higher or lower level of care. In the post-view questionnaire, clinicians were also asked if they intentionally planned to implement any of the following interventions after seeing the viewer: (1) changing any current medications, (2) ordering additional medications, (3) ordering additional diagnostic tests, (4) changing ventilation settings, (5) ordering additional labs, (6) physically reexamining this patient, (7) providing fluid bolus, or (8) providing a blood transfusion.

Statistical Methods

A participant's perceiving of the VS viewer's usefulness is represented by a vector consisting of the percentage of the 5 categories (strongly agree, agree, neutral, disagree, and strongly disagree) that he or she assigned to the question "the viewer enhanced my understanding of the patient's condition." We used the Ward method, a hierarchical clustering method, with Manhattan distance to group the participants based on their ratings to the question "the viewer enhanced my understanding of the patient's condition" [17,18]. Between those clusters, we compared the participants' opinion changes on the patients' conditions in 7 questions (Table 1) before and after using the viewer. The chi-square test was used to compare percentage differences.

Table 1. The number of opinion changes for 7 questions (Q1-Q7) before and after seeing the viewer, with respect to the 5 clustered user types.

Questions[a]	Total changes, n (%)	Unique participants, n (%)	C1, n	C2, n	C3, n	C4, n	C5, n	Like (C1 and C2), n (%)	Dislike (C3, C4, and C5), n (%)
Q1	129 (14.2)	16 (66.7)	46	31	10	42	0	77 (59.7)	52 (40.3)
Q2	112 (12.3)	15 (62.5)	38	34	8	32	0	72 (64.3)	40 (35.7)
Q3-6	145 (16)	18 (75)	58	54	3	30	0	112 (77.2)	33 (22.8)
Q7	92 (10.1)	17 (70.8)	20	32	9	31	0	52 (56.5)	40 (43.5)

[a]Please refer to Textbox 1 for the question.

XSL•FO
RenderX

Textbox 1. Questions.

- Q1: Having reviewed the last 24 hours of information during rounds and before and after seeing the 24-hour viewer, do they feel that in the past 24 hours the patient has shown evidence of (a) infection, (b) hemodynamic instability, (c) uncontrolled bleeding, or (d) respiratory deterioration?
- Q2: Over the past 24 hours, has the patient's condition (a) improved significantly, (b) improved slightly, (c) unchanged, (d) deteriorated slightly, or (e) deteriorated significantly?
- Q3: Can the patient be transferred to a lower level of care?
- Q4: Can the patient be transferred to a higher level of care?
- Q5: Does the patient have a traumatic brain injury?
- Q6: Did the patient have intracranial pressure problems in the past 24 hours?
- Q7: Due to the viewer, do they plan for any changes in interventions, including (a) changing any current medications, (b) ordering additional medications, (c) ordering additional diagnostic tests, (d) changing ventilation settings, (e) ordering additional labs, (f) physically reexamining this patient, (g) providing a fluid bolus, or (h) providing a blood transfusion?

Note: These are the questions referenced in Table 1.

Ethical Considerations

The study has been approved by the institutional review board of the University of Maryland School of Medicine (HP-00063086).

Results

Survey Collection

From February 2017 to June 2017, the survey team followed clinicians who agreed to take the surveys. A total of 908 surveys were collected from 24 participants with unbalanced proportions. Among the 908 rounds, 48 (5%) were patients who were newly admitted, and 860 (95%) were not. When asked if the VS viewer enhanced their understanding of the patient's condition, clinicians strongly agreed 45 (5%) times, agreed 514 (56.6%) times, were neutral 321 (35.4%) times, disagreed 26 (2.9%) times, and strongly disagreed 2 (0.2%) times. Figure 4 lists the total surveys each participant contributed and the proportions of ratings on whether the viewer enhanced his or her understanding of the patient's condition during a round.

Results show that physicians' clinical assessments and plans could be influenced by viewing the VS viewer for 1 minute or less, indicated by a "yes" answer to at least 1 of the 8 questions (Q7 in the survey). Of the 908 rounds, a total of 92 (10.1%) rounds had at least 1 "yes" as planning on some changes to the interventions. The most common change was (Q1) changing current medications (36/908, 4%). The next most common changes were (Q6) physically reexamining the patient (31/908, 3.4%), (Q2) ordering additional medications (20/908, 2.2%), and (Q7) providing a fluid bolus (20/908, 2.2%).

We used the Ward method with Manhattan distance to group the participants based on their ratings to the question "the viewer enhanced my understanding of the patient's condition" [17]. For example, 1 participant contributed 62 surveys and rated 2 "strongly agree," 22 "agree," 32 "neutral," 5 "disagree," and 1 "strongly disagree." The vector of percentages (0.03, 0.35, 0.52, 0.08, and 0.02) represents the overall rating that this participant had about the viewer. The 24 participants were clustered into 5 groups, as shown in Figure 5. The 5 groups correspond to the participants who are mostly in favor (C1) of the viewer to those least in favor (C5). There are 6 in C1, 6 in C2, 3 in C3, 7 in C4, and 2 in C5, which shows a very balanced grouping, with half of the participants in the C1 and C2 groups and the other half in the other 3 clusters. This shows that the sampled rounds were done by participants with almost similar proportions of different attitudes toward the viewer. In other words, the survey team sampled the rounds randomly enough so that the collected data were not biased by participants with certain preexisting feelings about the viewer.

XSL•FO
RenderX

Figure 4. Distribution of each participant's rating on if the viewer enhanced his or her understanding of the patient's condition during a round.

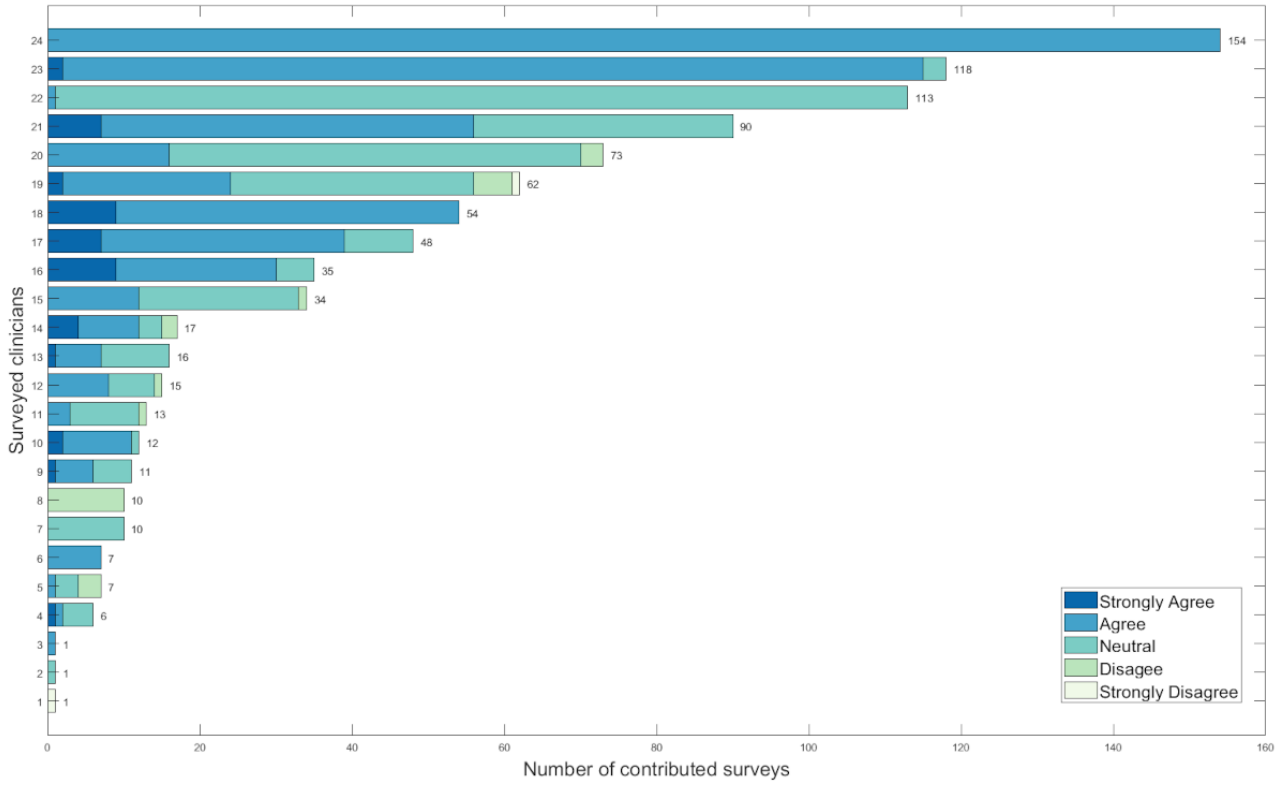
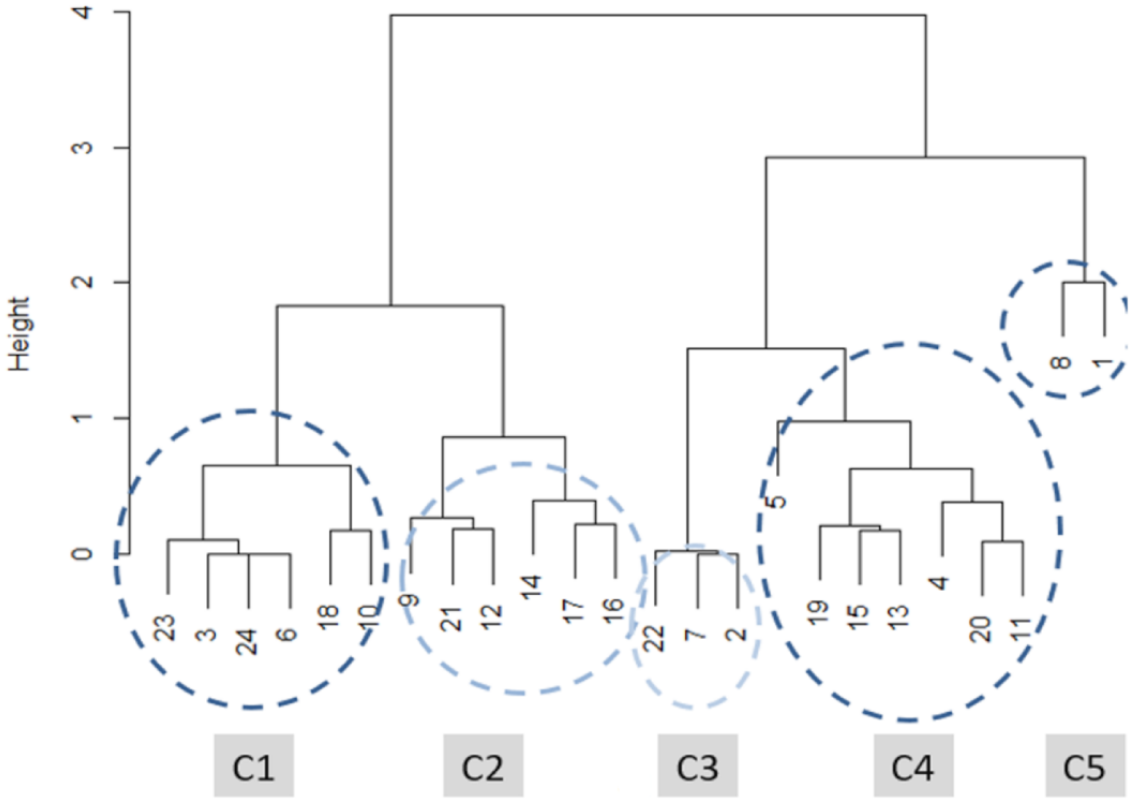


Figure 5. Clusters of participants with similar feelings about the viewer. In total, 24 unique participants are grouped into 5 groups, corresponding to "strongly favor," "favor," "neutral," "dislike," and "strongly dislike.".



Comparisons

We analyzed the opinion changes before and after seeing the viewer, regarding the patient's stability, status, and need for a higher or lower level of care. Instead of summarizing the total changes in opinions, we compared them with respect to the clusters of user types. The participants who were "neutral" (C3) or "strongly dislike" (C5) had low numbers of opinion changes

XSL•FO
RenderX

for all 7 questions. Those who were in clusters C1, C2, and C3 had more numbers of opinion changes (Table 1). For simplicity, we can further group the participants into 2 types: those who liked the VS viewer (C1 and C2) and those who disliked it (C3, C4, and C5). The clinicians who liked the VS viewer had a higher rate of changed opinions than those who disliked the VS viewer regarding Q1 to Q6 (Q1: 59.7% vs 40.3%, Q2: 64.3% vs 35.7%, and Q3-6: 77.2% vs 22.8%). When asked if they planned for any changes for interventions (Q7), there was no significant difference between the 2 major groups of clinicians (56.5% vs 43.5%, P=.10).

Discussion

Principal Results

With the development of sensor and computing technologies, vast amounts of high-quality, continuous electronic data, including VS, alarms, and clinical interventions, are collected at the bedside. Those data have the potential to provide an unprecedented view of dynamic physiologic responses to injury, illness, and treatments. Therefore, data gathered from bedsides could assist clinicians in care planning and decision support. However, massive amounts of data that are not well organized or presented still create a barrier for clinicians making full use of them in a busy resuscitation or intensive care environment. Bedside monitors often only display instantaneous readings or a short strip of recent physiologic VS for diagnosis. Clinicians need to rely on separate nursing charts, handwritten or electronic, to review a patient's developing conditions. The VS viewer, which automates physiological data by displaying clear color-coded trends, presents aggregated information from multiple data sources, provides at-a-glance summaries of clinical data, and assists with the prioritization of care for multiple patients.

The use of the VS viewer was subjectively assessed with 908 observations from clinicians working in ICUs at a high-volume level 1 trauma center. Clinicians generally perceived the use of the VS viewer favorably, as evidenced by survey data. The VS viewer was originally developed for the United States Air Force Critical Care Air Transport Teams [19,20]. Critical Care Air Transport Teams transport up to 3 patients who are critically ill in the back of the aircraft, allowing trauma surgeons to perform far-forward damage control surgery, knowing that these patients could be quickly transported rearward with full support. This rapid transport of complex patients with multisystem trauma, shock, burns, and respiratory failure who are in hemodynamic flux requires continual resuscitation, stabilization, advanced care, and life-saving interventions during air transport; however, currently available advanced ICU monitoring systems suitable for the needs of such patients were developed for use in stable, hospital-based settings, not in the crowded, noisy, vibrating, and sometimes frankly jolting environment of air evacuation or long-distance air transport. The noise levels, confined space, limited access to patients, vibration, and overall limited patient visibility make using a VS viewer advantageous in such a setting. Such technology can also be valuable in enhancing emergency medical personnel's decision-making for initial triage. While traditional VS are useful in guiding prehospital care and triage, they represent isolated points in time, and trends and fluctuations in vitals may not be apparent.

In this study, we set the clinical thresholds for colored warnings to be uniform across all beds. This was to make the user interface simplified and more consistent during a survey. Additionally, a set of predefined thresholds from a group of experienced clinicians could be a useful out-of-the-box feature when the VS viewer is deployed in the field. That said, the clinical thresholds could be personalized for each bed. For example, if the bedside monitor allows alarm threshold settings, such settings could be used as the colored warning thresholds in the VS viewer for each bed.

The VS viewer has expanded from ICUs to trauma resuscitation units, operating rooms, neuro ICUs, and pediatric ICUs at the University of Maryland Medical Center. In 2020, during the COVID-19 pandemic, it was deployed to monitor 150 beds in biocontaminated units to reduce the risk of infection and improve efficiency for clinicians in treating their patients.

Innovations

The VS viewer is a multipatient physiological monitor. To the best of our knowledge, we could not find any articles that describe a viewer system with a similar design. In a comprehensive review by Waller et al [5], a total of 17 information displays in ICU settings were designed for specific disease states or body systems, such as cerebral perfusion monitoring for individual patients or monitoring for arterial blood gas trends. The novel user interface presented in this study was designed with the aim of conveying information more efficiently to ICU clinicians in a noisy, confined, and busy environment. It uses color-coded warnings to indicate a patient's status and highlight data that needs attention. The side panel provides a peek at the physiological status of other patients, which can help clinicians keep an eye on other patients even if their attention is focused on a single patient. It uses advanced web front-end techniques to hide large quantities of data behind simple line charts and reveal them when needed.

Clinical Impact

The use of the VS viewer can have several possible influences on clinical assessment and plans. It can help clinicians quickly recognize critical changes in the patient's physiologic status and provide early interventions to prevent further deterioration. The VS viewer can potentially improve patient outcomes by providing clinicians with a concise overview of key information, reducing cognitive load and errors, and improving compliance with evidence-based safety guidelines [12]. It may also help to improve communication efficiency within the ICU team by providing easy access to a shared platform of patient longitudinal data. It can reduce the workload of the ICU team by automating routine tasks such as extracting data from nursing charts.

To prioritize care in high-volume ICUs, intensive care clinicians must be able to rapidly identify physiological events and the need for intervention. The VS viewer can help organize a large amount of data in a busy, noisy ICU environment where close monitoring of patients who are critically ill is essential to detect potentially harmful physiological trends. The presentation of data with temporal, color-coded patterns, and the ability of the

XSL•FO
RenderX

VS viewer to provide at-a-glance data for entire units is advantageous for clinicians working in high-volume ICUs.

The color-coded patterns may reduce the "alarm fatigue" issue in noisy ICUs. The noise burden is common in modern physiologic monitoring systems and has been recognized as a critical patient safety concern in the hospital care setting [21-23]. In noisy environments, such as ICUs, helicopter transportation, or aeromedical evacuation, loud and continuous alarms could reduce their specificity in getting clinicians' responses. Another issue with audible alarms is that they are transient and cannot be replayed once they are gone. While the visual alert patterns could show the longitudinal patterns of physiologic change.

Related Work

The VS viewer with organized and easy-access information could be part of the effort to build the smart ICU or the tele-ICU. The concepts of smart ICU and tele-ICU aim to maximize the use of bedside clinical expertise in assessing and treating patients by providing integrated monitoring and actionable information [24-26]. A survey study of 86 ICU staff in a German university hospital summarized that health providers expect ICU monitoring could be improved by reducing false alarms, using wireless sensors and mobile devices, preparing for the use of AI, and enhancing the digital literacy of ICU staff [27,28]. The VS viewer could be used in both centralized and decentralized architectures of tele-ICU for extending coverage and facilitating patient transfer between hospitals because of its flexible configuration of grouping ICU beds virtually [29]. By making essential clinical information available remotely, the VS viewer allows clinicians to provide care plans when on-site support is infeasible or limited [30,31]. It may potentially reduce exposure to contagious diseases and, hence, increase patient safety.

With continuous physiologic data and other clinical information, the VS viewer has the ability to process real-time data into predictive algorithms, which is also desired for tele-ICU [30]. Beyond being a plain displayer, the VS viewer could embed risk-prediction algorithms that use continuous VS as inputs and may promote more efficient interventions to reduce ICU risk [31]. For example, ICU mortality prediction [32,33], secondary insults after severe brain traumatic injury [34], needs for transfusion [35,36], and neurologic decline in the ICU [37] are reported to have good predictive performances by using variables derived from continuous VS. We have also shown that using risk scores calculated from continuously measured VS, patients requiring endovascular resuscitative interventions can be identified with high accuracy [38]. Moreover, the VS viewer could serve as a platform for predictive model diagnosis by providing clinicians with explainable artificial intelligence [39]. With patient VS data, we can use the Shapley Additive exPlanations algorithm to calculate each variable's contribution to the prediction result [40]. Therefore, the clinicians would know not only the prediction but also the contribution of each variable to the prediction. Such information may help clinicians make more personalized care plans.

Limitations

There are limitations to this work that are worth noting. We collected data from a large number of ICU clinicians compared to trauma team clinicians. Trauma team clinicians are surgeons responsible for the same patient throughout the entire length of stay, regardless of the acuity of the patient. ICU clinicians are intensivists and are only responsible for patients in the ICU. Hence, disparities between both groups of clinicians are inevitable, as each group has different clinical perspectives and patient workloads. As occurs in nearly all survey work, response rates and receptiveness to the surveys varied. Some clinicians were more amenable to being surveyed compared to others. In the collected forms, there were more surveys from some clinicians than from others. To reduce this potential bias, we clustered the participants based on their overall rating on each round, from which we estimated each participant's a priori attitude toward using this viewer. The results show that there was a balanced "favoring" and "non-favoring" of using this viewer.

We only evaluated the viewer based on clinicians' satisfaction and efficiency (potential changes in interventions before and after seeing the viewer). In future studies, randomized controlled trials can be designed to analyze the viewer's impact on patients' outcomes and safety [12].

Conclusions

We designed, implemented, and evaluated an automated physiologic data organizer and visualization platform. It provides at-a-glance summaries and assists with prioritizing care for multiple patients. The VS viewer demonstrates a method to assemble large quantities of data from multiple sources and represents trends in each patient's condition with simple color codes, greatly improving situational awareness. It has the potential to be used in en route care, hospitals with multiple branches, and understaffed hospitals in remote areas. The survey shows that organized physiologic data and visual assessment could assist clinicians in recognizing changes in patient status and prioritizing care.

Acknowledgments

This work was partially funded by USAF 8650-14-2-6D19 and W81XWH-17-C-0034.

The authors thank the VS viewer study group: Colin Mackenzie, MD, Sarah Wade, BS, Lauren Hartsky, BS, Thomas Grissom, MD, Napoleon Roux, MD, Doug Floccare, MD, Ashton Engdahl, BS, Scott Murray, BS, Rosemary Kozar, MD, Sarah Murthi, MD, Laura Buchanan, MD, Neal Reynolds, MD, Gunjan Parikh, MD, Megan Brenner, MD, without whose help the data collection for this study would not have been possible.

XSL•FO
RenderX

Conflicts of Interest

PH, DS, Colin Mackenzie (part of the VS viewer study group), TS, and SY have US Patent Application 17/676,657 filed on February 21, 2022, titled “Method and Apparatus for Monitoring Collection of Physiological Patient Data.”

Multimedia Appendix 1
Surveyed thresholds for heart rate, systolic and diastolic blood pressure, blood oxygen saturation, and temperature.
[DOCX File , 20 KB - humanfactors_v11i1e46030_app1.docx]

Multimedia Appendix 2
VS Viewer color coding threshold values.
[DOCX File , 18 KB - humanfactors_v11i1e46030_app2.docx]

References

1. Endsley MR. Designing for Situation Awareness: An Approach to User-Centered Design, 2nd Edition. United Kingdom: CRC Press; 2016.
2. Yang S, Njoku M, Mackenzie CF. 'Big data' approaches to trauma outcome prediction and autonomous resuscitation. Br J Hosp Med (Lond) 2014 Nov;75(11):637-641. [doi: 10.12968/hmed.2014.75.11.637] [Medline: 25383434]
3. Drew BJ, Harris P, Zègre-Hemsey JK, Mammone T, Schindler D, Salas-Boni R, et al. Insights into the problem of alarm fatigue with physiologic monitor devices: a comprehensive observational study of consecutive intensive care unit patients. PLoS One 2014;9(10):e110274 [FREE Full text] [doi: 10.1371/journal.pone.0110274] [Medline: 25338067]
4. Sebastian K, Sari V, Loy LY, Zhang F, Zhang Z, Feng M. Multi-signal Visualization of Physiology (MVP): A novel visualization dashboard for physiological monitoring of Traumatic Brain Injury patients. : IEEE; 2012 Presented at: 2012 Annual International Conference of the IEEE Engineering in Medicine and Biology Society; 28 August 2012-01 September 2012; San Diego, CA, USA p. 2000-2003 URL: https://ieeexplore.ieee.org/document/6346349 [doi: 10.1109/embc.2012.6346349]
5. Waller RG, Wright MC, Segall N, Nesbitt P, Reese T, Borbolla D, et al. Novel displays of patient information in critical care settings: a systematic review. J Am Med Inform Assoc 2019 May 01;26(5):479-489 [FREE Full text] [doi: 10.1093/jamia/ocy193] [Medline: 30865769]
6. Excel Medical Electronics, LLC. BedMasterEx Operator's Manual. Jupiter, FL: Excel Medical Electronics, LLC; 2013.
7. Hu PF, Yang S, Li HC, Stansbury LG, Yang F, Hagegeorge G, et al. Reliable collection of real-time patient physiologic data from less reliable networks: a "Monitor of Monitors" system (MoMs). J Med Syst 2017 Jan;41(1):3. [doi: 10.1007/s10916-016-0648-5] [Medline: 27817131]
8. Dziadzko MA, Herasevich V, Sen A, Pickering BW, Knight AMA, Moreno Franco P. User perception and experience of the introduction of a novel critical care patient viewer in the ICU setting. Int J Med Inform 2016 Apr;88:86-91. [doi: 10.1016/j.ijmedinf.2016.01.011] [Medline: 26878767]
9. Pickering BW, Herasevich V, Ahmed A, Gajic O. Novel Representation of Clinical Information in the ICU: Developing User Interfaces which Reduce Information Overload. Appl Clin Inform 2010;1(2):116-131 [FREE Full text] [doi: 10.4338/ACI-2009-12-CR-0027] [Medline: 23616831]
10. Ahmed A, Chandra S, Herasevich V, Gajic O, Pickering BW. The effect of two different electronic health record user interfaces on intensive care provider task load, errors of cognition, and performance. Crit Care Med 2011 Jul;39(7):1626-1634. [doi: 10.1097/CCM.0b013e31821858a0] [Medline: 21478739]
11. Roger SP, Bruce RM. Software Engineering: A Practitioner's Approach. United States: McGraw-Hill Education; 2015.
12. Khairat SS, Dukkipati A, Lauria HA, Bice T, Travers D, Carson SS. The Impact of Visualization Dashboards on Quality of Care and Clinician Satisfaction: Integrative Literature Review. JMIR Hum Factors 2018 May 31;5(2):e22 [FREE Full text] [doi: 10.2196/humanfactors.9328] [Medline: 29853440]
13. Garrett JJ. Ajax: a new approach to web applications. AdaptivePath.com. 2005. URL: https://designftw.mit.edu/lectures/apis/ajax_adaptive_path.pdf [accessed 2022-09-28]
14. King RW, Plewa MC, Buderer NM, Knotts FB. Shock index as a marker for significant injury in trauma patients. Acad Emerg Med 1996 Nov;3(11):1041-1045 [FREE Full text] [doi: 10.1111/j.1553-2712.1996.tb03351.x] [Medline: 8922013]
15. Kahraman S, Hu P, Stein DM, Stansbury LG, Dutton RP, Xiao Y, et al. Dynamic three-dimensional scoring of cerebral perfusion pressure and intracranial pressure provides a brain trauma index that predicts outcome in patients with severe traumatic brain injury. J Trauma 2011 Mar;70(3):547-553. [doi: 10.1097/TA.0b013e31820c768a] [Medline: 21610341]
16. Laughery KR, Wogalter MS. Designing effective warnings. Rev Hum Factors Ergonomics 2016 Nov 03;2(1):241-271. [doi: 10.1177/1557234x0600200109]
17. Murtagh F, Legendre P. Ward’s hierarchical agglomerative clustering method: which algorithms implement Ward’s criterion? J Classif 2014 Oct 18;31(3):274-295. [doi: 10.1007/s00357-014-9161-z]
18. Strauss T, von Maltitz MJ. Generalising Ward's method for use with Manhattan distances. PLoS One 2017;12(1):e0168288 [FREE Full text] [doi: 10.1371/journal.pone.0168288] [Medline: 28085891]

XSL•FO
RenderX

19. Lairet J, King J, Vojta L, Beninati W. Short-term outcomes of US Air Force Critical Care Air Transport Team (CCATT) patients evacuated from a combat setting. Prehosp Emerg Care 2013;17(4):486-490. [doi: 10.3109/10903127.2013.811564] [Medline: 23865806]
20. Ingalls N, Zonies D, Bailey JA, Martin KD, Iddins BO, Carlton PK, et al. A review of the first 10 years of critical care aeromedical transport during operation iraqi freedom and operation enduring freedom: the importance of evacuation timing. JAMA Surg 2014 Aug;149(8):807-813. [doi: 10.1001/jamasurg.2014.621] [Medline: 25074327]
21. Cvach M. Monitor alarm fatigue: an integrative review. Biomed Instrum Technol 2012;46(4):268-277. [doi: 10.2345/0899-8205-46.4.268] [Medline: 22839984]
22. Joint Commission. Medical device alarm safety in hospitals. Sentinel Event Alert 2013 Apr 08(50):1-3. [Medline: 23767076]
23. Oliveira AECD, Machado AB, Santos EDD, Almeida ÉBD. Alarm fatigue and the implications for patient safety. Rev Bras Enferm 2018;71(6):3035-3040 [FREE Full text] [doi: 10.1590/0034-7167-2017-0481] [Medline: 30517409]
24. Halpern NA. Innovative designs for the smart ICU. Chest 2014 Mar;145(3):646-658. [doi: 10.1378/chest.13-0004] [Medline: 27845639]
25. Halpern NA. Innovative designs for the smart ICU: part 3: advanced ICU informatics. Chest 2014 Apr;145(4):903-912. [doi: 10.1378/chest.13-0005] [Medline: 24687712]
26. Avdalovic MV, Marcin JP. When will telemedicine appear in the ICU? J Intensive Care Med 2019 Apr;34(4):271-276. [doi: 10.1177/0885066618775956] [Medline: 29756527]
27. Poncette AS, Spies C, Mosch L, Schieler M, Weber-Carstens S, Krampe H, et al. Clinical requirements of future patient monitoring in the intensive care unit: qualitative study. JMIR Med Inform 2019 Apr 30;7(2):e13064 [FREE Full text] [doi: 10.2196/13064] [Medline: 31038467]
28. Poncette AS, Mosch L, Spies C, Schmieding M, Schiefenhövel F, Krampe H, et al. Improvements in patient monitoring in the intensive care unit: survey study. J Med Internet Res 2020 Jun 19;22(6):e19091 [FREE Full text] [doi: 10.2196/19091] [Medline: 32459655]
29. Guinemer C, Boeker M, Fürstenau D, Poncette AS, Weiss B, Mörgeli R, et al. Telemedicine in intensive care units: scoping review. J Med Internet Res 2021 Nov 03;23(11):e32264 [FREE Full text] [doi: 10.2196/32264] [Medline: 34730547]
30. Hassan E. Tele-ICU and patient safety considerations. Crit Care Nurs Q 2018;41(1):47-59. [doi: 10.1097/CNQ.0000000000000185] [Medline: 29210766]
31. Udeh C, Udeh B, Rahman N, Canfield C, Campbell J, Hata JS. Telemedicine/virtual ICU: where are we and where are we going? Methodist Debakey Cardiovasc J 2018;14(2):126-133 [FREE Full text] [doi: 10.14797/mdcj-14-2-126] [Medline: 29977469]
32. Baker S, Xiang W, Atkinson I. Continuous and automatic mortality risk prediction using vital signs in the intensive care unit: a hybrid neural network approach. Sci Rep 2020 Dec 04;10(1):21282. [doi: 10.1038/s41598-020-78184-7] [Medline: 33277530]
33. Kristinsson Æ, Gu Y, Rasmussen SM, Mølgaard J, Haahr-Raunkjær C, Meyhoff CS, et al. Prediction of serious outcomes based on continuous vital sign monitoring of high-risk patients. Comput Biol Med 2022 Aug;147:105559 [FREE Full text] [doi: 10.1016/j.compbiomed.2022.105559] [Medline: 35635901]
34. Bonds BW, Yang S, Hu PF, Kalpakis K, Stansbury LG, Scalea TM, et al. Predicting secondary insults after severe traumatic brain injury. J Trauma Acute Care Surg 2015 Jul;79(1):85-90; discussion 90. [doi: 10.1097/TA.0000000000000698] [Medline: 26091319]
35. Parimi N, Hu PF, Mackenzie CF, Yang S, Bartlett ST, Scalea TM, et al. Automated continuous vital signs predict use of uncrossed matched blood and massive transfusion following trauma. J Trauma Acute Care Surg 2016 Jun;80(6):897-906. [doi: 10.1097/TA.0000000000001047] [Medline: 27027555]
36. Yang S, Mackenzie CF, Rock P, Lin C, Floccare D, Scalea T, et al. Comparison of massive and emergency transfusion prediction scoring systems after trauma with a new Bleeding Risk Index score applied in-flight. J Trauma Acute Care Surg 2021 Feb 01;90(2):268-273. [doi: 10.1097/TA.0000000000003031] [Medline: 33502145]
37. Podell J, Yang S, Miller S, Felix R, Tripathi H, Parikh G, et al. Rapid prediction of secondary neurologic decline after traumatic brain injury: a data analytic approach. Sci Rep 2023 Jan 09;13(1):403 [FREE Full text] [doi: 10.1038/s41598-022-26318-4] [Medline: 36624110]
38. Zeineddin A, Hu P, Yang S, Floccare D, Lin CY, Scalea TM, et al. Prehospital continuous vital signs predict need for resuscitative endovascular balloon occlusion of the aorta and resuscitative thoracotomy prehospital continuous vital signs predict resuscitative endovascular balloon occlusion of the aorta. J Trauma Acute Care Surg 2021 Nov 01;91(5):798-802. [doi: 10.1097/TA.0000000000003171] [Medline: 33797486]
39. Maul J, Straub J. Assessment of the Use of Patient Vital Sign Data for Preventing Misidentification and Medical Errors. Healthcare (Basel) 2022 Dec 02;10(12):1-87 [FREE Full text] [doi: 10.3390/healthcare10122440] [Medline: 36553964]
40. Lundberg S, Lee SI. A unified approach to interpreting model predictions. ArXiv ;30:1-10 Preprint posted online on 25 Nov 2017. [FREE Full text]

XSL•FO
RenderX

Abbreviations

HR: heart rate
ICP: intracranial pressure
ICU: intensive care unit
SI: shock index
VS: vital signs

Edited by A Kushniruk; submitted 27.02.23; peer-reviewed by AS Poncette, J Walsh, K Gupta; comments to author 14.06.23; revised version received 03.11.23; accepted 20.11.23; published 05.01.24.

Please cite as:
Yang S, Galvagno S, Badjatia N, Stein D, Teeter W, Scalea T, Shackelford S, Fang R, Miller C, Hu P, VS viewer study group
A Novel Continuous Real-Time Vital Signs Viewer for Intensive Care Units: Design and Evaluation Study
JMIR Hum Factors 2024;11:e46030
URL: https://humanfactors.jmir.org/2024/1/e46030
doi:10.2196/46030
PMID:38180791

©Shiming Yang, Samuel Galvagno, Neeraj Badjatia, Deborah Stein, William Teeter, Thomas Scalea, Stacy Shackelford, Raymond Fang, Catriona Miller, Peter Hu, VS viewer study group. Originally published in JMIR Human Factors (https://humanfactors.jmir.org), 05.01.2024. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on https://humanfactors.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Characterizing and Comparing Adverse Drug Events Documented in 2 Spontaneous Reporting Systems in the Lower Mainland of British Columbia, Canada: Retrospective Observational Study

Erica Y Lau[1,2*], PhD; Amber Cragg[2], MSc; Serena S Small[2], MA; Katherine Butcher[3], PharmB; Corinne M Hohl[1*], MSc, MD

[1]Department of Emergency Medicine, University of British Columbia, Vancouver, BC, Canada

[2]Centre for Clinical Epidemiology and Evaluation, Vancouver Coastal Health Research Centre, Vancouver, BC, Canada

[3]Pharmaceutical Science, Vancouver General Hospital, Vancouver Coastal Health, Vancouver, BC, Canada

[*]these authors contributed equally

Corresponding Author:
Erica Y Lau, PhD
Department of Emergency Medicine
University of British Columbia
7th Floor, 828 West 10th Avenue Research Pavilion
Vancouver, BC, V5Z 1M9
Canada
Phone: 1 6048754111 ext 68926
Email: erica.lau@ubc.ca

Abstract

Background: Robust adverse drug event (ADE) reporting systems are crucial to monitor and identify drug safety signals, but the quantity and type of ADEs captured may vary by system characteristics.

Objective: We compared ADEs reported in 2 different reporting systems in the same jurisdictions, the Patient Safety and Learning System–Adverse Drug Reaction (PSLS-ADR) and ActionADE, to understand report variation.

Methods: This retrospective observational study analyzed reports entered into PSLS-ADR and ActionADE systems between December 1, 2019, and December 31, 2022. We conducted a comprehensive analysis including all events from both reporting systems to examine coverage and usage and understand the types of events captured in both systems. We calculated descriptive statistics for reporting facility type, patient demographics, serious events, and most reported drugs. We conducted a subanalysis focused on adverse drug reactions to enable direct comparisons between systems in terms of the volume and events reported. We stratified results by reporting system.

Results: We performed the comprehensive analysis on 3248 ADE reports, of which 12.4% (375/3035) were reported in PSLS-ADR and 87.6% (2660/3035) were reported in ActionADE. Distribution of all events and serious events varied slightly between the 2 systems. Iohexol, gadobutrol, and empagliflozin were the most common culprit drugs (173/375, 46.2%) in PSLS-ADR, while hydrochlorothiazide, apixaban, and ramipril (308/2660, 11.6%) were common in ActionADE. We included 2728 reports in the subanalysis of adverse drug reactions, of which 12.9% (353/2728) were reported in PSLS-ADR and 86.4% (2357/2728) were reported in ActionADE. ActionADE captured 4- to 6-fold more comparable events than PSLS-ADR over this study's period.

Conclusions: User-friendly and robust reporting systems are vital for pharmacovigilance and patient safety. This study highlights substantial differences in ADE data that were generated by different reporting systems. Understanding system factors that lead to varying reporting patterns can enhance ADE monitoring and should be taken into account when evaluating drug safety signals.

(JMIR Hum Factors 2024;11:e52495) doi:10.2196/52495

KEYWORDS

adverse drug event reporting systems; side effect; side effects; drug; drugs; pharmacy; pharmacology; pharmacotherapy; pharmaceutic; pharmaceutics; pharmaceuticals; pharmaceutical; medication; medications; patient safety; health information

XSL•FO
RenderX

technology; pharmacovigilance; adverse; safety; HIT; information system; information systems; reporting; descriptive statistics; monitoring

Introduction

Over 2 million Canadians visit an emergency department every year because of an adverse drug event (ADE), an unintended and harmful event related to medication use [1,2]. ADEs incur over 700,000 hospital admissions, and cost over CAD $1 billion (USD $7.48 million) in annual health care expenditures across Canada [2,3]. The importance of addressing this issue cannot be overstated: the World Health Organization (WHO) has identified the prevention of ADEs as an urgent global public health priority [4].

In response to this pressing concern, Canada implemented regulations outlined in the Protecting Canadians from Unsafe Drugs Act (Vanessa's Law) which came into full effect on December 16, 2019. This federal legislation mandates prompt reporting of serious adverse drug reactions (ADRs; a subtype of ADEs) and medical device incidents from hospitals to Health Canada within 30 days of documentation [5]. These regulations serve as a safeguard to protect patients and improve drug surveillance.

Postmarketing pharmacovigilance is crucial in the detection, assessment, and prevention of ADEs under real-world conditions [6,7]. Among the various methods used, spontaneous reporting stands out as one of the most widely adopted approaches in pharmacovigilance [8]. When patients or health professionals spontaneously report ADEs, drug safety monitoring agencies evaluate and integrate these reports into databases, enabling ongoing identification of safety signals [7,8]. This method of surveillance captures data from a broad population and allows us to detect drug safety signals that may not have been identified in the randomized trials used for drug licensing and monitor rare ADEs to medications [9].

It is important to recognize, however, that there is considerable variation in ADE reporting systems worldwide in terms of their design, data fields, terminologies [10], and implementations, which may impact the volume and type of ADEs reported [11]. Variation in design also leads to a lack of standardization of reports, which can in turn prohibit interoperability or effective exchange of ADE reports between systems and may prevent comparisons of ADE events, rates, and risk factors across systems [10].

Despite this variation, the diversity of systems may also be a strength. Each system has the potential to complement others, enhancing the overall quantity and quality of ADE data, if variation in design leads to variation in reporting behaviors or the types of reports that can be entered [12]. To leverage this untapped potential, we need to better understand and compare the events collected through diverse reporting systems [13]. Understanding similarities and differences between systems will enable researchers and drug safety monitoring agencies to more effectively use existing data for accurate signal detection, especially for new or rare ADEs, and prioritize the investigation of drug safety signals. This knowledge will also aid stakeholders in optimizing the design and implementation of new reporting systems to enhance ADE data collection and drug safety surveillance and better align systems with their intended purpose [10].

Health Canada, the regulatory authority for postmarketing pharmacovigilance in Canada, oversees the Canada Vigilance Program, collecting reports of suspected ADEs since 1965. Health professionals and consumers can voluntarily submit reports through various channels, including a web-based platform, phone, fax, or mail. Hospitals are required to submit written reports within 30 days, and Health Canada allows them flexibility by permitting the use of existing systems and processes to meet reporting requirements. With Health Canada's approval, hospitals may use a third party, such as a regional health authority or other reporting programs, to submit reports [14].

The province of British Columbia (BC) currently uses 2 approved spontaneous reporting systems that enable hospitals to comply with Vanessa's Law mandates: the BC Patient Safety and Learning System–Adverse Drug Reaction (PSLS-ADR) reporting form and ActionADE. Briefly, PSLS-ADR was developed and implemented as the first province-wide, web-based platform and supports hospitals in meeting the mandatory reporting requirements [14]. ActionADE, implemented later in the timeline, is a research-driven, web-based app that aims to prevent unintentional redispensation of harmful medications by facilitating the sharing of ADE information between providers across health care settings. ADE reporting occurs as a byproduct of enabling safer care provision (Multimedia Appendix 1) [15].

These 2 systems enable a comparison of the quality and quantity of ADE data generated using 2 different designs. Our objective is to describe and compare the ADEs that health care providers documented using PSLS-ADR and ActionADE during the first 3 years following the implementation of Vanessa's Law.

Methods

ADE Reporting Systems

About PSLS-ADR

BC Patient Safety and Learning System (PSLS) is an initiative of the BC Patient Safety Task Force, developed in collaboration with all 6 provincial health authorities and the Health Care Protection Program, which is part of the Risk Management Branch of the Ministry of Finance that insures BC hospitals [16]. BC PSLS is a web-based safety event reporting and management information system designed to support the identification, investigation, and analysis of safety and risk-related events, including safety hazards, near misses, and adverse events [17]. The system underwent a pilot phase in 2007 and was subsequently implemented province-wide in 2008. BC PSLS has been instrumental in promoting patient safety within the health care system in BC [16].

XSL•FO
RenderX

In response to the introduction of Vanessa's Law and in collaboration with Health Canada, BC PSLS launched PSLS-ADR as a new add-on to the existing system in 2014 and released an updated version in 2019 [18,19]. PSLS-ADR is accessible to health care facilities in all health authorities across BC, including acute care hospitals, long-term care facilities, and outpatient clinics. Authorized health care professionals with access to the secure health authority network, including employees, medical staff, paramedics, contractors, students, and volunteers, can submit reports to PSLS-ADR [15]. Once a report is submitted, the system notifies the medication safety officer in the respective health authority to review and respond to the event [20]. The health authorities send eligible reports to Health Canada for Vanessa's Law reporting requirements. Reports are not made available to care providers and not integrated into the electronic medical record. They are only generated for the purposes of pharmacosurveillance (Multimedia Appendix 2).

The PSLS-ADR data fields are based largely on the Canada Vigilance Adverse Reaction Reporting Form, with additional questions enabling medication safety officers, pharmacy representatives, and others to follow-up with reporters or patients, if necessary [20]. The PSLS-ADR reporting form contains 26 required data fields that collect information about the patient, the adverse reaction (eg, seriousness), the suspected health products (types, name, route used, therapy dates, and treatments), and the reporters.

About ActionADE

Previous studies found that 32.5% of ADE cases observed in emergency departments are repeat events [21], often occurring due to the unintentional represcription or redispensation of the same or a same-class medication as one that previously caused harm [22]. This recurrence is attributed to the lack of effective means to communicate and integrate ADE information into clinical workflows. ActionADE, a research-driven initiative, was developed to address this communication gap [23,24].

In collaboration with the Ministry of Health, Vancouver Coastal Health, a technology partner, and health professional organizations and clinicians, our research team developed and piloted ActionADE between 2016 and 2019 using participatory design principles and data standards that were evaluated and subsequently pilot tested to optimize the system's usability [10,11,15,24-28]. In 2020, we began the implementation of ActionADE in 1 hospital (Vancouver General Hospital) and then expanded its use to 6 hospitals operated by Vancouver Coastal Health and Providence Health Care as part of a research initiative. Although providers were encouraged to use ActionADE, they maintain complete autonomy in choosing between the PSLS-ADR and ActionADE systems to meet their needs.

ActionADE is a web-based app that allows providers to document and communicate ADE information, bidirectionally through its integration (or linkage) with BC's central drug database (PharmaNet). ActionADE was accessible to a subset of care providers with an eligible prescriber identification number issued by their respective regulatory college (ie, physicians, pharmacists, and nurse practitioners) [29]. Eligible clinicians submit reports to ActionADE from a designated health authority network, and the data are shared with clinicians within the patient's circle of care via PharmaNet and used to create safety alerts when community pharmacists attempt to redispense culprit or same-class medications. ActionADE complements the PSLS-ADR system by automating ADE reporting to Health Canada (Multimedia Appendix 3).

The ActionADE data fields were developed based on a systematic review of ADE reporting systems worldwide and participatory action research with clinician end users and are compatible with Health Canada's Canada Vigilance Adverse Reaction Reporting Form [10,11,15,27,30]. As ActionADE is integrated with PharmaNet, several fields auto-populate based on the patient's personal health number, including patient's personal and demographic information (ie, name, date of birth, and sex), reporter's information (ie, name, role, and site), and patient's 14-month medication dispensation history. To create a new report, the system auto-populates the patient's information and medication dispensation history, as well as the reporter's information. ActionADE contains 5 required data fields that collect information about the suspect drugs, which is auto generated based on the medication dispensation history or added manually, the ADE type, and details of the event (eg, symptoms or diagnosis, outcome, and certainty; Multimedia Appendix 4).

Study Design

In this retrospective observational study, we analyzed reports documented in PSLS-ADR and ActionADE entered by providers at health care facilities operated by the Vancouver Coastal Health Authority (excluding Providence Health Care, as PSLS-ADR data were unavailable from those facilities) in BC, Canada, between December 1, 2019, to December 31, 2022.

For PSLS-ADR, we included reports documented by authorized health care professionals (eg, employees, medical staff, paramedics, contractors, students, and volunteers) from >120 health care facilities across the province, including hospital, urgent and primary care, long-term care facilities, and community health centers' clinics. For ActionADE, we included reports documented by eligible clinicians from 4 hospitals where ActionADE is implemented: Lions Gate Hospital, Richmond Hospital, UBC (University of British Columbia) Hospital, and Vancouver General Hospital (Multimedia Appendix 5) [31].

We divided this study's period into 4 phases: baseline period, when all hospitals across BC only used PSLS-ADR (December 2019 to February 2020); year 1 (March 2020 to November 2020), when 1 hospital (Vancouver General Hospital) had the option to use ActionADE for piloting purposes while all other sites in BC exclusively used PSLS-ADR; and year 2 (December 2020 to November 2021) and year 3 (December 2021 to December 2022), when the 4 hospitals had the option to use PSLS-ADR or ActionADE and all other sites in BC exclusively used PSLS-ADR (Multimedia Appendix 6).

Data Sources

For this study, we requested ADE reports from PSLS-ADR from the BC PSLS central office and retrieved reports documented in ActionADE during the same period from the ActionADE database. We obtained information about hospital

XSL•FO
RenderX

characteristics through the Information Access and Privacy Services at Provincial Health Services Authority, including number of beds, population served, and the number of emergency department visits per year.

Data Extraction

To allow for direct comparisons between the 2 systems, we combined similar variables wherever possible. A clinical pharmacist classified all free-text drug entries from the PSLS-ADR reports into the equivalent generic drug name that would be present if the same report were entered into ActionADE based on the provincial formulary. We translated continuous age from ActionADE into the age categories in PSLS-ADR. We combined information across platforms to produce combined variables for report date, patient demographics (age group and sex), types of ADE (ADRs and nonadverse drug reactions), ADE outcomes (death, emergency visit, hospitalized or hospital extended, life threatening, worsened preexisting condition, permanent disability and fetal defect, other, and unknown) [27], and reporter information (role and facility).

Statistical Analysis

Comprehensive Analysis

First, we performed a comprehensive statistical analysis to provide a global view of coverage, usage, and the types of information captured by both reporting systems. We included all events from both reporting systems, excluding reports related to user errors (eg, duplicate reports), refuted allergies, and reports with incomplete data. We calculated descriptive statistics (eg, means and SDs or frequency and percentages) for the following variables: total number and types of the reporting facilities (hospital vs nonhospital), patient's age group, patient's sex, roles of reporters, proportion of serious events, and the 10 most reported culprit drugs for all events and serious events. We defined serious events based on the Health Canada's definition. This definition includes ADEs that require in-patient hospitalization or prolongation of existing hospitalization, cause congenital malformation, result in persistent or significant disability or incapacity, are life-threatening, or result in death [14].

ADR Analysis

To allow for direct comparisons between the 2 systems, we then conducted a subsample analysis that only included ADR reports (a subtype of ADE) that met Health Canada's definition and that could have been reported in both systems. According to Health Canada, ADRs encompass harmful and unintended responses to a health product, including any undesirable patient effects suspected to be associated with health product use. This definition includes unintended effects, health product abuse, overdoses, interactions (including drug-drug and drug-food interactions), and unusual lack of therapeutic efficacy, all of which are considered reportable adverse reactions [14,32,33]. We included eligible ADR reports from both reporting systems from sites where both systems were available, excluding reports related to user errors (eg, duplicate reports), refuted allergies, and reports with incomplete data. We calculated descriptive statistics for the following variables: patient's age group, patient's sex, proportion of serious events, the 10 most reported culprit drugs, and mean monthly counts of all events and serious events during each phase of this study period, stratified by reporting system. We conducted all analyses using SAS statistical software (version 9.4; SAS Institute).

Ethical Considerations

The UBC (University of British Columbia) clinical research ethics board approved of this research (H18-01332 and H22-00312) and provided a waiver for obtaining informed consent as this study meets the Tri-Council Policy Statement minimal risk criteria.

Results

Comprehensive Analysis

We extracted 3248 reports from both reporting systems. After removing 213 reports related to refuted allergies, erroneous reports, and reports with incomplete data, the analytic cohort for the comprehensive analysis comprised 3035 unique ADEs reported in either system (Figure 1). Of these, 12.4% (375/3035) were entered in PSLS-ADR and 87.6% (2660/3035) were reported in ActionADE.

Approximately 50% of the events occurred in male patients in both PSLS-ADR (178/375) and ActionADE (1285/3035). The highest proportion of events were from patients aged 45-64 years (32.8%, 123/375) in PSLS-ADR and aged 75-84 years (25.3%, 674/2660) in ActionADE. In total, 12 facilities (5 hospitals and 7 nonhospital facilities) entered reports in PSLS-ADR. The primary reporters in PSLS-ADR were medical imaging staff or technicians (170/375, 45.3%) and pharmacists (174/375, 46.4%). Of the 4 hospitals that entered reports in ActionADE, pharmacists were the reporter for 92.1% (2451/3035) of the events. The proportion of serious events was 36% (135/377) in PSLS-ADR and 28.2% (749/3035) in ActionADE (Table 1).

In PSLS, the most common culprit drugs were iohexol, gadobutrol, and empagliflozin, accounting for 46.2% (173/375) of all events. Empagliflozin, ibuprofen, and iohexol represented 11.8% (16/135) of serious events (Tables 2 and 3). Iohexol and gadobutrol are both contrast agents used for diagnostic imaging, whereas empagliflozin is an oral medication primarily prescribed for managing type 2 diabetes mellitus and ibuprofen is an oral, over-the-counter nonsteroidal anti-inflammatory drug used to relieve pain, reduce inflammation, and alleviate fever.

In ActionADE, the most common culprit drugs were hydrochlorothiazide, ramipril, and apixaban, which accounted for 10.5% (356/3391) of all events; hydrochlorothiazide, empagliflozin, and apixaban represented 11.1% (105/951) of serious events (Tables 2 and 3). Hydrochlorothiazide and ramipril are commonly prescribed for hypertension. Apixaban, an oral anticoagulant, is primarily used for stroke prevention in patients with atrial fibrillation and for treatment and prevention of venous thromboembolism.

XSL•FO
RenderX

Figure 1. Flow diagram. ADE: adverse drug event; ADR: adverse drug reaction; PSLS-ADR: Patient Safety and Learning System–Adverse Drug Reaction.

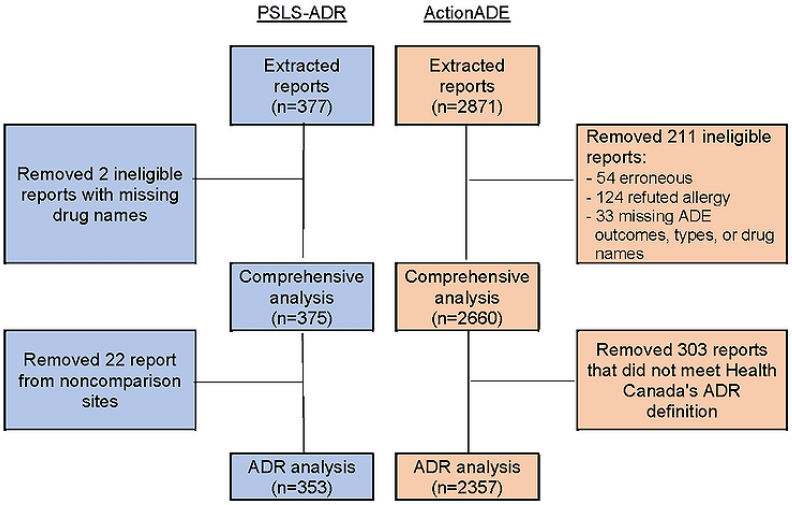


Table 1. Descriptive statistics of all events included in the comprehensive analysis by reporting system[a].

Characteristics	PSLS-ADR[b] (n=375), n (%)	ActionADE[c] (n=2660), n (%)
Type of reporting facilitates		
Hospitals	5 (41.6)	4 (100)
Nonhospitals	7 (58.4)	0 (0)
Patient age group (y)		
<1-19	11 (2.9)	22 (0.8)
20-44	83 (22.1)	299 (11.2)
45-64	123 (32.8)	523 (19.7)
65-74	67 (17.9)	544 (20.5)
75-84	52 (13.9)	674 (25.3)
>84	39 (10.4)	598 (22.5)
Patient sex		
Male	178 (47.5)	1285 (48.3)
Role of reporter		
Physicians	Suppressed[d]	204 (7.7)
Nurses	27 (7.2)	__[e]
Medical imaging staff or technologists	170 (45.3)	__[e]
Nurse practitioners	Suppressed	5 (0.2)
Pharmacists	174 (46.4)	2451 (92.1)
Others	Suppressed	__[e]
Proportion of serious events[f]	135 (36.0)	749 (28.2)

[a]The comprehensive analysis included all events from both reporting systems excluding reports related to errors, refuted allergy, and incomplete data on study variables.

[b]PSLS-ADR: Patient Safety and Learning System–Adverse Drug Reaction.

[c]ADE: adverse drug event.

[d]Cell sizes <5 are suppressed.

[e]These personnel are not eligible to report in ActionADE.

[f]Serious events are those with an outcome of fetal defect, permanent disability, hospitalization, extended hospitalization, life threatening, or death.

XSL•FO
RenderX

Table 2. Most frequently reported culprit drugs for all events in the comprehensive analysis by reporting systems.

System and drug	n (%)
PSLS-ADR[a] (n=375)	
Iohexol	154 (41.1)
Gadobutrol	12 (3.2)
Empagliflozin	7 (1.9)
Rivaroxaban	7 (1.9)
Furosemide	6 (1.6)
Nivolumab	6 (1.6)
Ramipril	6 (1.6)
Unknown generic drug	6 (1.6)
Acetylsalicylic acid	5 (1.3)
Ibuprofen	5 (1.3)
ActionADE[b] (n=2660)	
Hydrochlorothiazide	113 (4.2)
Apixaban	103 (3.9)
Ramipril	92 (3.5)
Acetylsalicylic acid	88 (3.3)
Warfarin	88 (3.3)
Rivaroxaban	79 (3)
Furosemide	77 (2.9)
Empagliflozin	63 (2.4)
Metformin HCL[c]	52 (2)
Spironolactone	50 (1.9)

[a]PSLS-ADR: Patient Safety and Learning System–Adverse Drug Reaction.
[b]ADE: adverse drug event.
[c]HCL: hydrochloride.

XSL•FO
RenderX

Table 3. Most frequently reported culprit drugs for serious events[a] in the comprehensive analysis by reporting systems.

System and drug	n (%)
PSLS-ADR[b] (n=135)	
Empagliflozin	6 (4.4)
Ibuprofen	5 (3.7)
Iohexol	5 (3.7)
Nivolumab	5 (3.7)
Acetylsalicylic acid	Suppressed[c]
Glyburide	Suppressed
Rivaroxaban	Suppressed
Allopurinol	Suppressed
Amlodipine besylate	Suppressed
Apixaban	Suppressed
ActionADE[d] (n=749)	
Hydrochlorothiazide	43 (5.7)
Empagliflozin	26 (3.5)
Apixaban	24 (3.2)
Furosemide	20 (2.7)
Acetylsalicylic acid	18 (2.4)
Rivaroxaban	18 (2.4)
Candesartan cilexetil	16 (2.1)
Ramipril	16 (2.1)
Chlorthalidone	16 (2.1)
Spironolactone	15 (2)

[a]Serious events are those with an outcome of fetal defect, permanent disability, hospitalization, extended hospitalization, life threatening, or death.

[b]PSLS-ADR: Patient Safety and Learning System–Adverse Drug Reaction.

[c]Cell sizes <5 are suppressed.

[d]ADE: adverse drug event.

ADR Analysis

We included a total of 2728 reports that met Health Canada's definition of an ADR from facilities that had the option of using either reporting system during this study's period (Figure 1) [32,33]. Of the included reports, 12.9% (353/2728) were entered in PSLS-ADR, while the majority (2357/2728, 86.4%) were reported in ActionADE.

The distribution of ADR reports by patient sex, age, primary reporters and proportion of serious events for both systems were similar to the comprehensive analysis (Table 4). However, each reporting system revealed distinct patterns of reporting. In PSLS-ADR, iohexol, gadobutrol, and empagliflozin accounted for 44.8% (168/353) of all events, while empagliflozin, ibuprofen, and nivolumab represented 12.1% (16/133) of serious events. In ActionADE, hydrochlorothiazide, ramipril, and apixaban accounted for 12% (284/2357) of all events. Furthermore, hydrochlorothiazide, empagliflozin, and apixaban represented 13.4% (88/671) of serious events (Tables 5 and 6).

A direct comparison in events reportable through both the PSLS-ADR and ActionADE systems revealed an increase in event reporting, including serious events, following the implementation of ActionADE (Figures 2 and 3). Baseline measurements indicate that the mean monthly counts of all events and serious events across sites were 2.9 (95% CI 2.2 to 3.6) and 1.7 (95% CI 0.8 to 2.5), respectively. In period 3, the mean monthly counts of all events and serious events across sites escalated to 27.2 (95% CI 20.4 to 34.0) and 7.0 (95% CI 4.9 to 9.2), respectively, reflecting a 9- and 4-fold increase over time. Furthermore, the mean monthly counts of all events and serious events during this study's period within the ActionADE system were 6- and 4-fold greater than that of PSLS-ADR.

XSL•FO
RenderX

Table 4. Descriptive statistics of events meeting Health Canada's ADR[a] definition[b] across common reporting sites[c] by reporting system.

Characteristics	PSLS-ADR[d] (n=353), n (%)	ActionADE[e] (n=2357), n (%)
Patient age group (y)		
<1-19	11 (3.1)	18 (0.8)
20-44	77 (21.8)	239 (10.1)
45-64	114 (32.3)	450 (19.1)
65-74	64 (18.1)	494 (21)
75-84	49 (13.9)	606 (25.7)
>84	38 (10.8)	550 (23.3)
Patient sex		
Male	173 (49)	1114 (47.3)
Role of reporter		
Physicians	Suppressed[f]	161 (6.8)
Nurses	22 (6.2)	—[g]
Medical imaging staff or technologists[h]	155 (43.9)	—
Nurse practitioners	Suppressed	5 (0.2)
Pharmacists	173 (49)	2190 (92.9)
Others[h]	Suppressed	—
Proportion of serious events[i]	133 (37.7)	671 (28.5)

[a]ADR: adverse drug reactions.

[b]According to Health Canada adverse drug reaction includes unintended effects, health product abuse, overdoses, interactions (including drug-drug and drug-food interactions), and unusual lack of therapeutic efficacy.

[c]Common reporting sites included Vancouver General, University of British Columbia, Lions Gate, and Richmond Hospitals.

[d]PSLS-ADR: Patient Safety and Learning System–Adverse Drug Reaction.

[e]ADE: adverse drug event.

[f]Cell sizes <5 are suppressed.

[g]Not available.

[h]These personnel are not eligible to report in ActionADE.

[i]Serious events are those with an outcome of fetal defect, permanent disability, hospitalization, extended hospitalization, life threatening, or death.

XSL•FO
RenderX

Table 5. Most frequently reported culprit drugs for all events meeting Health Canada's ADR[a] definitions[b] across common reporting sites[c] by reporting system and severity.

System and drug	n (%)
PSLS-ADR[d] (n=353)	
Iohexol	139 (39.4)
Gadobutrol	12 (3.4)
Empagliflozin	7 (2)
Rivaroxaban	7 (2)
Furosemide	6 (1.7)
Nivolumab	6 (1.7)
Ramipril	6 (1.7)
Acetylsalicylic acid	5 (1.4)
Ibuprofen	5 (1.4)
Indapamide	5 (1.4)
ActionADE[e] (n=2357)	
Hydrochlorothiazide	109 (4.6)
Ramipril	88 (3.7)
Apixaban	87 (3.7)
Acetylsalicylic acid	75 (3.2)
Warfarin	74 (3.1)
Rivaroxaban	71 (3)
Empagliflozin	60 (2.5)
Furosemide	56 (2.4)
Metformin HCL[f]	44 (1.9)
Ibuprofen	43 (1.8)

[a]ADR: adverse drug reaction.

[b]According to Health Canada adverse drug reaction includes unintended effects, health product abuse, overdoses, interactions (including drug-drug and drug-food interactions), and unusual lack of therapeutic efficacy.

[c]Common reporting sites included Vancouver General, University of British Columbia, Lions Gate, and Richmond Hospitals.

[d]PSLS-ADR: Patient Safety and Learning System–Adverse Drug Reaction.

[e]ADE: adverse drug event.

[f]HCL: hydrochloride.

XSL•FO
RenderX

Table 6. Most frequently reported culprit drugs for serious events[a] meeting Health Canada's ADR[b] definitions[c] across common reporting sites[d] by reporting system and severity.

System and drug	n (%)
PSLS-ADR[e] (n=133)	
Empagliflozin	6 (4.5)
Ibuprofen	5 (3.8)
Nivolumab	5 (3.8)
Acetylsalicylic acid	Suppressed[f]
Glyburide	Suppressed
Rivaroxaban	Suppressed
Allopurinol	Suppressed
Amlodipine besylate	Suppressed
Apixaban	Suppressed
Clopidogrel bisulfate	Suppressed
ActionADE[g] (n=671)	
Hydrochlorothiazide	43 (6.7)
Empagliflozin	25 (3.7)
Apixaban	20 (3)
Acetylsalicylic acid	19 (2.8)
Chlorthalidone	16 (2.4)
Ramipril	14 (2.1)
Rivaroxaban	14 (2.1)
Ibuprofen	13 (1.9)
Warfarin	13 (1.9)
Candesartan cilexetil	12 (1.8)

[a]Common reporting sites included Vancouver General, University of British Columbia, Lions Gate, and Richmond Hospitals.

[b]ADR: adverse drug reaction.

[c]According to Health Canada adverse drug reaction includes unintended effects, health product abuse, overdoses, interactions (including drug-drug and drug-food interactions), and unusual lack of therapeutic efficacy.

[d]Serious events are those with an outcome of fetal defect, permanent disability, hospitalization, extended hospitalization, life threatening, or death.

[e]PSLS-ADR: Patient Safety and Learning System–Adverse Drug Reaction.

[f]Cell sizes <5 are suppressed.

[g]ADE: adverse drug event.

XSL•FO
RenderX

Figure 2. Mean monthly counts of all events meeting Health Canada's ADR definitions across common reporting sites during this study's period. ADR: adverse drug reaction; PSLS-ADR: Patient Safety and Learning System–Adverse Drug Reaction.

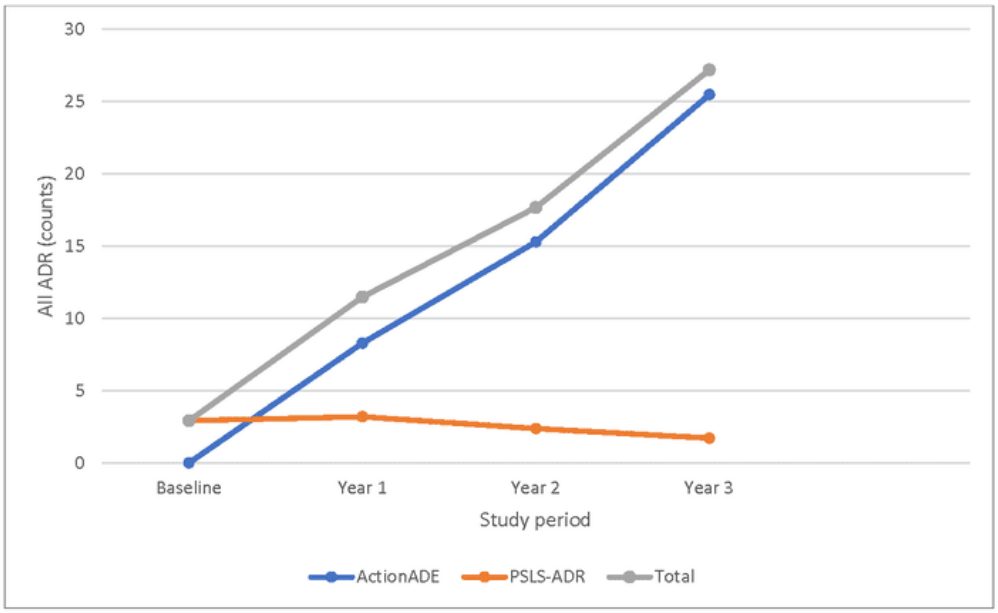
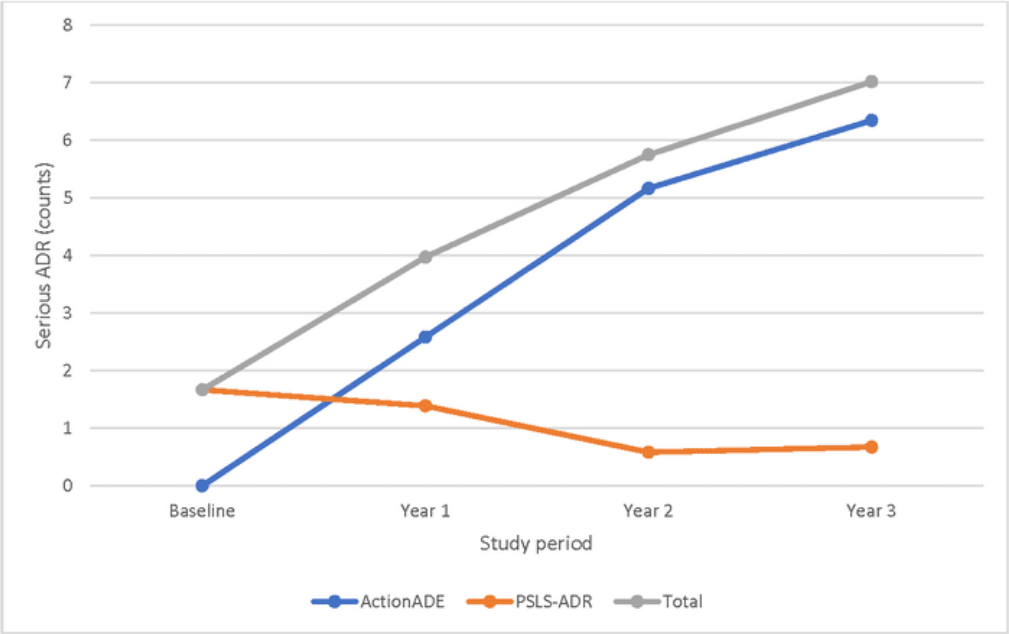


Figure 3. Mean monthly counts of serious events meeting Health Canada's ADR definitions across common reporting sites during this study's period. ADR: adverse drug reaction; PSLS-ADR: Patient Safety and Learning System–Adverse Drug Reaction.



Discussion

Principal Findings

Our study aimed to describe and compare ADEs reported using 2 distinct reporting systems that were developed and implemented in different ways. Both PSLS-ADR and ActionADE are currently in use in BC in the first 3 years following the implementation of Vanessa's Law. We observed differences in reports between the 2 systems regarding their coverage, usage, and the type of ADE data captured.

PSLS-ADR had broader coverage, collecting data from various health care facilities including community health centers, vaccination clinics, and outpatient clinics. Its user base was more diverse including physicians, nurses, medical imaging staff or technologists, nurse practitioners, pharmacists, and other professionals. In contrast, ActionADE coverage was limited to ADEs identified in patients presenting to 4 participating hospitals, with clinical pharmacists as its primary user. The broader coverage of PSLS-ADR can be attributed to its established position as a provincial safety event reporting platform; its accessibility to a broader range of health professions; and a federal mandate for hospitals to stimulate reporting using health authority wide communication efforts including email blasts, information on health authority websites, and presentations to provider groups. Leveraging the insights gained from PSLS-ADR, our research team is actively collaborating with key stakeholders to broaden ActionADE's

XSL•FO
RenderX

app. The Vancouver Coastal Health Authority, where ActionADE is presently in use, has recently endorsed it as a standard practice for ADE reporting in new care settings, including long-term care homes, in-patient wards, and community clinics.

Although PSLS-ADR exhibited broader coverage, ActionADE demonstrated higher usage. Our comparative analysis revealed that the average monthly counts of all events and serious events in ActionADE were 6 and 4 times higher, respectively, than in the PSLS-ADR system. Several factors might contribute to these discrepancies in reporting rates. First, PSLS-ADR was designed solely for Vanessa's Law compliance, with reports forwarded to Health Canada for surveillance purposes. ActionADE, on the other hand, serves the dual purpose of functioning as both a clinical communication tool and a means of complying with Vanessa's Law, thus improving patient safety [15]. Reports entered into ActionADE are used to generate preventive alerts in community pharmacies when pharmacists attempt redispensation of a drug that has previously caused the patient harm, which have demonstrated preliminary effectiveness [34]. The potential impact of reporting in ActionADE on patient safety is likely a motivating factor for providers to report ADEs [35]. Furthermore, ActionADE has a proactive implementation support mechanism, which has been shown to be instrumental in enhancing providers' adoption of the reporting platform [35]. Finally, ActionADE used participatory design principles to optimize its design to facilitate use by end users and is integrated with PharmaNet to enable prepopulation of fields to allow reporters to generate reports ≤2 minutes, whereas PSLS-ADR users noted that reports can take 20 minutes to complete [34].

The 2 systems captured adverse events to different culprit drugs. This can be attributed to the more limited accessibility of ActionADE. The most reported drugs in PSLS-ADR were iohexol and gadobutrol, and correspondingly, medical imaging staff or technologists made up a significant proportion of reporters. This suggests that the current workflow for ADR reporting of radiopharmaceuticals is designated to medical imaging staff or technologists. Imaging staff or technologists were unable to use ActionADE at the time of this study due to PharmaNet legislation, which requires that users have prescriber ID restricting use to physicians, pharmacists, and nurse practitioners. This restriction has resulted in fewer radiopharmaceutical ADRs to be reported, as pharmacists generally do not work in radiology departments.

ActionADE frequently captured hydrochlorothiazide-related events, while only a few of such events were captured in PSLS-ADR. Among the ADRs associated with hydrochlorothiazide, electrolyte disturbances, and acute kidney injury were found to be the most common [34], involving multiple additional contributing factors. The specific functionality offered by ActionADE, such as the ability to specify the provider's certainty that the patient's presentation and the option to update or refute events based on new information or alternative diagnoses, likely played a role in encouraging clinicians to report these more complex events [11,15,27,30].

Ibuprofen was the second most commonly reported culprit drug related to serious events in PSLS-ADR, but it barely made the top 10 in ActionADE. This discrepancy may be due to the over-the-counter status of ibuprofen, which means patients can access the medication without a prescription and bypass communication about ADEs from ActionADE that is built into the prescription dispensation process.

While our study primarily focused on comparing these 2 systems, it is crucial to view these findings in the broader context of ADE reporting. Despite these disparities, both systems play vital roles in contributing to patient safety by capturing valuable information on ADEs. PSLS-ADR is an effective means of capturing radiopharmaceutical-related ADEs by imaging staff and technicians who are not trained in taking medication histories or ADE assessments, while ActionADE is more effective for pharmaceutical-related ADEs by clinical pharmacists that are reported and communicated on a patient-level to improve safety. These systems work in a complementary manner, catering to different areas of the health care system and capturing unique data and thus offering a more comprehensive picture of ADEs. For example, a common signal between the 2 systems might indicate a more serious issue for a specific drug irrespective of context (eg, empagliflozin). These findings suggest the need for careful attention to the design and implementation of these systems to ensure they effectively serve their intended users and context of use and ensure data resulting from each system are interpreted correctly by end users. The absence of reporting of one type of event may reflect design, implementation, or user characteristics rather than the absence of these events.

Limitations

To our knowledge, this is the first study to directly compare 2 ADE reporting systems operating within the same jurisdiction. While the results of our study provided valuable insights into the differences between these systems, it is important to acknowledge several limitations that warrant consideration in interpreting the results. First, our study sample was confined to 2 reporting systems, which may not fully encapsulate the diversity of all systems employed across health care settings globally. As a result, the findings may not be generalizable to other reporting systems. Second, our data set was limited to facilities that used PSLS-ADR or ActionADE for reporting. This reduces the generalizability of our findings to the wider array of health care facilities in BC or nationally. It is plausible that unaccounted-for variations in data and reporting practices among facilities not deploying these 2 systems could exist. Third, our study may be susceptible to unmeasured and uncontrolled confounding variables. For example, the level of organizational emphasis on ADE reporting, differences in implementation, available resources, and providers' perceptions could have affected the usage and coverage of the 2 systems under study. This variability might have further influenced the nature of ADE information reported. Fourth, the relatively small number of drugs resulting in ADEs prevented us from conducting a robust quantitative comparison of these events. Furthermore, the data we used were a snapshot in time and may not reflect changes in reporting systems or health care facilities that have occurred since then. Lastly, we consciously chose not

XSL•FO
RenderX

to draw comparisons with other studies examining the frequently reported culprit drugs from spontaneous reporting systems in other jurisdictions. This decision stems from the recognition that the diversity in ADE reports—both in terms of numbers and types—is intricately tied to factors such as system design, geography, population characteristics, drug exposures, and the medical system itself. To facilitate meaningful comparisons across studies, a more robust surveillance system is needed.

Conclusions

Understanding the differences between reporting systems can inform future systems design and improvement, including changes to user training and implementation, and inform the use of forthcoming data and procurement decisions for reporting systems. Further research could explore how to integrate the strengths of both systems, potentially leading to more comprehensive safety data to facilitate drug and patient safety and inform pharmacoepidemiologic studies. Continuous evaluation and improvement are essential considering the significant role these systems play to improve our health systems.

Acknowledgments

The ActionADE platform is funded by the Canadian Institutes of Health Research, with support from the Ministry of Health British Columbia, Vancouver Coastal Health, the Lower Mainland Consolidated Pharmacies, and the BC SUPPORT Unit. CMH is supported by the Michael Smith Foundation for Health Research Health Professional-Investigator award. EYL is supported by the Michael Smith Health Research BC Research Trainee Award (GR018656). The funder has no role in the design, conduct, analysis, interpretation of data, or the drafting of scientific papers.

Authors' Contributions

All authors contributed to this study's conception and design. EYL requested access and analyzed the data and wrote the first draft of this paper. AC, CMH, and SSS contributed to the refinement of the data analysis. All authors contributed to the interpretation of the findings and commented on previous paper versions. All authors read and approved the submitted paper and have agreed to be personally accountable for their contribution.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Characteristics of (Patient Safety and Learning System–Adverse Drug Reaction) PSLS-ADR and ActionADE.
[DOCX File , 23 KB - humanfactors_v11i1e52495_app1.docx]

Multimedia Appendix 2
Screenshot of Patient Safety and Learning System–Adverse Drug Reaction (PSLS-ADR).
[DOCX File , 282 KB - humanfactors_v11i1e52495_app2.docx]

Multimedia Appendix 3
Screenshot of ActionADE.
[DOCX File , 527 KB - humanfactors_v11i1e52495_app3.docx]

Multimedia Appendix 4
Data fields included in Patient Safety and Learning System–Adverse Drug Reaction (PSLS-ADR) and ActionADE.
[DOCX File , 36 KB - humanfactors_v11i1e52495_app4.docx]

Multimedia Appendix 5
Characteristics of sites that had options to use Patient Safety and Learning System–Adverse Drug Reaction (PSLS-ADR) or ActionADE systems.
[DOCX File , 30 KB - humanfactors_v11i1e52495_app5.docx]

Multimedia Appendix 6
Descriptions of the 4-phase study period.
[DOCX File , 159 KB - humanfactors_v11i1e52495_app6.docx]

References

XSL•FO
RenderX

1. Hohl CM, Dankoff J, Colacone A, Afilalo M. Polypharmacy, adverse drug-related events, and potential adverse drug interactions in elderly patients presenting to an emergency department. Ann Emerg Med 2001;38(6):666-671. [doi: 10.1067/mem.2001.119456] [Medline: 11719747]
2. Zed PJ, Abu-Laban RB, Balen RM, Loewen PS, Hohl CM, Brubacher JR, et al. Incidence, severity and preventability of medication-related visits to the emergency department: a prospective study. CMAJ 2008;178(12):1563-1569 [FREE Full text] [doi: 10.1503/cmaj.071594] [Medline: 18519904]
3. NACRS emergency department visits and lengths of stay. Canadian Institute for Health Information. Ottawa, Ontario: Canadian Institute for Health Information; 2018. URL: https://www.cihi.ca/en/nacrs-emergency-department-visits-and-lengths-of-stay [accessed 2023-12-22]
4. Sheikh A, Dhingra-Kumar N, Kelley E, Kieny MP, Donaldson LJ. The third global patient safety challenge: tackling medication-related harm. Bull World Health Organ 2017;95(8):546-546A [FREE Full text] [doi: 10.2471/BLT.17.198002] [Medline: 28804162]
5. Herder M, Gibson E, Graham J, Lexchin J, Mintzes B. Regulating prescription drugs for patient safety: does Bill C-17 go far enough? CMAJ 2014;186(8):E287-E292 [FREE Full text] [doi: 10.1503/cmaj.131850] [Medline: 24616135]
6. Raj N, Fernandes S, Charyulu NR, Dubey A, Hebbar GS. Postmarket surveillance: a review on key aspects and measures on the effective functioning in the context of the United Kingdom and Canada. Ther Adv Drug Saf 2019;10:2042098619865413 [FREE Full text] [doi: 10.1177/2042098619865413] [Medline: 31384423]
7. The importance of pharmacovigilance: safety monitoring of medical products. World Health Organization. 2002. URL: https://www.who.int/publications/i/item/10665-42493 [accessed 2022-01-23]
8. Alomar M, Tawfiq AM, Hassan N, Palaian S. Post marketing surveillance of suspected adverse drug reactions through spontaneous reporting: current status, challenges and the future. Ther Adv Drug Saf 2020;11:2042098620938595 [FREE Full text] [doi: 10.1177/2042098620938595] [Medline: 32843958]
9. Hazell L, Shakir SAW. Under-reporting of adverse drug reactions : a systematic review. Drug Saf 2006;29(5):385-396 [FREE Full text] [doi: 10.2165/00002018-200629050-00003] [Medline: 16689555]
10. Bailey C, Peddie D, Wickham ME, Badke K, Small SS, Doyle-Waters MM, et al. Adverse drug event reporting systems: a systematic review. Br J Clin Pharmacol 2016;82(1):17-29 [FREE Full text] [doi: 10.1111/bcp.12944] [Medline: 27016266]
11. Hohl CM, Small SS, Peddie D, Badke K, Bailey C, Balka E. Why clinicians don't report adverse drug events: qualitative study. JMIR Public Health Surveill 2018;4(1):e21 [FREE Full text] [doi: 10.2196/publichealth.9282] [Medline: 29487041]
12. Fletcher AP. Spontaneous adverse drug reaction reporting vs event monitoring: a comparison. J R Soc Med 1991;84(6):341-344 [FREE Full text] [doi: 10.1177/014107689108400612] [Medline: 2061900]
13. Lynn RM, Riding K, McIntosh N. The use of electronic reporting to aid surveillance of ADRs in children: a proof of concept study. Arch Dis Child 2010;95(4):262-265 [FREE Full text] [doi: 10.1136/adc.2009.158766] [Medline: 20335236]
14. Mandatory Reporting of Serious Adverse Drug Reactions and Medical Device Incidents by Hospitals: Guidance Document. Ottawa, ON: Health Canada; 2019.
15. Hohl C, Lexchin JR, Balka E. Can reporting of adverse drug reactions create safer systems while improving health data? CMAJ 2015;187(11):789-790 [FREE Full text] [doi: 10.1503/cmaj.150057] [Medline: 25941261]
16. British Columbia patient safety & learning system annual report. BC Patient Safety & Learning System Central Office. 2008. URL: https://bcpslscentral.ca/wp-content/uploads/2014/04/pslsannualreport1011.pdf [accessed 2023-12-22]
17. Cochrane D, Taylor A, Miller G, Hait V, Matsui I, Bharadwaj M, et al. Establishing a provincial patient safety and learning system: pilot project results and lessons learned. Healthc Q 2009;12:147-153 [FREE Full text] [doi: 10.12927/hcq.2009.20717] [Medline: 19667793]
18. Online ADR reporting is coming!. British Columbia Patient Safety & Learning System Central. 2014. URL: https://bcpslscentral.ca/online-adr-reporting-is-coming/#:~:text=Most%20of%20us%20will%20take,taking%20the%20medication%20as%20intended [accessed 2023-12-22]
19. Vanessa's law (ADR & MDI reporting) fact sheet. British Columbia Patient Safety & Learning System Central. URL: https://medicalstaff.islandhealth.ca/sites/default/files/2019-11/ADR%20and%20MDI%20Reporting%20Fact%20Sheet%20and%20FAQ_0.pdf [accessed 2023-12-22]
20. Online ADR reporting is coming!. British Columbia Patient Safety & Learning System. 2014. URL: https://bcpslscentral.ca/online-adr-reporting-is-coming/ [accessed 2023-05-23]
21. Woo SA, Cragg A, Wickham ME, Peddie D, Balka E, Scheuermeyer F, et al. Methods for evaluating adverse drug event preventability in emergency department patients. BMC Med Res Methodol 2018;18(1):160 [FREE Full text] [doi: 10.1186/s12874-018-0617-4] [Medline: 30514232]
22. van der Linden CMJ, Jansen PAF, van Marum RJ, Grouls RJE, Korsten EHM, Egberts ACG. Recurrence of adverse drug reactions following inappropriate re-prescription: better documentation, availability of information and monitoring are needed. Drug Saf 2010;33(7):535-538 [FREE Full text] [doi: 10.2165/11532350-000000000-00000] [Medline: 20553055]
23. Hau JP, Brasher PMA, Cragg A, Small S, Wickham M, Hohl CM. Using ActionADE to create information continuity to reduce re-exposures to harmful medications: study protocol for a randomized controlled trial. Trials 2021;22(1):119 [FREE Full text] [doi: 10.1186/s13063-021-05061-7] [Medline: 33546752]

XSL•FO
RenderX

24. Hohl CM, Woo SA, Cragg A, Wickham ME, Ackerley C, Scheuermeyer F, et al. Repeat adverse drug events associated with outpatient medications: a descriptive analysis of 3 observational studies in British Columbia, Canada. CMAJ Open 2019;7(3):E446-E453 [FREE Full text] [doi: 10.9778/cmajo.20180190] [Medline: 31320328]
25. Chruscicki A, Badke K, Peddie D, Small S, Balka E, Hohl CM. Pilot-testing an adverse drug event reporting form prior to its implementation in an electronic health record. Springerplus 2016;5(1):1764 [FREE Full text] [doi: 10.1186/s40064-016-3382-z] [Medline: 27795906]
26. Hohl CM, Badke K, Zhao A, Wickham ME, Woo SA, Sivilotti MLA, et al. Prospective validation of clinical criteria to identify emergency department patients at high risk for adverse drug events. Acad Emerg Med 2018;25(9):1015-1026 [FREE Full text] [doi: 10.1111/acem.13407] [Medline: 29517818]
27. Peddie D, Small SS, Badke K, Bailey C, Balka E, Hohl CM. Adverse drug event reporting from clinical care: mixed-methods analysis for a minimum required dataset. JMIR Med Inform 2018;6(2):e10248 [FREE Full text] [doi: 10.2196/10248] [Medline: 29954724]
28. Chan E, Small SS, Wickham ME, Cheng V, Balka E, Hohl CM. The utility of different data standards to document adverse drug event symptoms and diagnoses: mixed methods study. J Med Internet Res 2021;23(12):e27188 [FREE Full text] [doi: 10.2196/27188] [Medline: 34890351]
29. PharmaCare policy manual (unified). Appendix C-practitioner ID reference codes. British Columbia Ministry of Health. 2023. URL: https://www2.gov.bc.ca/gov/content/health/practitioner-professional-resources/pharmacare/pharmacare-publications/pharmacare-policy-manual-unified/practitioner-id-reference-codes [accessed 2023-05-22]
30. Peddie D, Small S, Wickham M, Bailey C, Hohl C, Balka E. Designing novel health ICTs to support work, not generate it: five principles. Stud Health Technol Inform 2017;234:262-268. [Medline: 28186052]
31. How's BC doing so far with Vanessa's law? British Columbia Patient Safety & Learning System. 2012. URL: http://bcpslscentral.ca./ [accessed 2023-12-22]
32. Module 1: overview of Vanessa's law and reporting requirements. Health Canada. 2020. URL: https://www.canada.ca/en/health-canada/services/drugs-health-products/medeffect-canada/adverse-reaction-reporting/mandatory-hospital-reporting/education/module-1.html [accessed 2020-01-31]
33. Serious adverse drug reaction reporting form for hospitals. Health Canada. 2020. URL: https://www.canada.ca/content/dam/hc-sc/documents/services/drugs-health-products/medeffect-canada/adverse-reaction-reporting/mandatory-reporting-hospital-drug-eng.pdf [accessed 2023-07-13]
34. Cragg A, Small SS, Lau E, Rowe A, Lau A, Butcher K, et al. Sharing adverse drug event reports between hospitals and community pharmacists to inform re-dispensing: an analysis of reports and process outcomes. Drug Saf 2023;46(11):1161-1172 [FREE Full text] [doi: 10.1007/s40264-023-01348-7] [Medline: 37783974]
35. Lau EY, Small SS, Butcher K, Cragg A, Loh GW, Shalansky S, et al. An external facilitation intervention to increase uptake of an adverse drug event reporting intervention. Front Health Serv 2023;3:1106586 [FREE Full text] [doi: 10.3389/frhs.2023.1106586] [Medline: 37332530]

Abbreviations

ADE: adverse drug event
ADR: adverse drug reaction
BC: British Columbia
PSLS-ADR: Patient Safety and Learning System–Adverse Drug Reaction
PSLS: Patient Safety and Learning System
UBC: University of British Columbia
WHO: World Health Organization

Edited by K Adapa; submitted 05.09.23; peer-reviewed by R Lim, A Austin; comments to author 28.11.23; revised version received 06.12.23; accepted 07.12.23; published 18.01.24.

Please cite as:
Lau EY, Cragg A, Small SS, Butcher K, Hohl CM
Characterizing and Comparing Adverse Drug Events Documented in 2 Spontaneous Reporting Systems in the Lower Mainland of British Columbia, Canada: Retrospective Observational Study
JMIR Hum Factors 2024;11:e52495
URL: https://humanfactors.jmir.org/2024/1/e52495
doi:10.2196/52495
PMID:38236629

XSL•FO
RenderX

©Erica Y Lau, Amber Cragg, Serena S Small, Katherine Butcher, Corinne M Hohl. Originally published in JMIR Human Factors (https://humanfactors.jmir.org), 18.01.2024. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on https://humanfactors.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Usability and Evaluation of a Health Information System in the Emergency Department: Mixed Methods Study

Christina Østervang[1,2], RN, MScN, PhD; Charlotte Myhre Jensen[2,3], RN, MEd, PhD; Elisabeth Coyne[2,4], RN, PhD; Karin B Dieperink[2,5,6], RN, MCN, PhD; Annmarie Lassen[1,2], DMSci, PhD

[1]Department of Emergency Medicine, Odense University Hospital, Odense, Denmark
[2]Department of Clinical Research, University of Southern Denmark, Odense, Denmark
[3]Department of Orthopedic Surgery and Traumatology, Odense University Hospital, Odense, Denmark
[4]School of Nursing and Midwifery, Griffith University, Brisbane, Australia
[5]Department of Oncology, Odense University Hospital, Odense, Denmark
[6]Family Focused Healthcare Research Center (FACE), University of Southern Denmark, Odense, Denmark

Corresponding Author:
Christina Østervang, RN, MScN, PhD
Department of Emergency Medicine, Odense University Hospital
Kløvervænget 25
Odense, 5000
Denmark
Phone: 45 21757009
Email: christina.oestervang@rsyd.dk

Abstract

Background: A lack of information during an emergency visit leads to the experience of powerlessness for patients and their family members, who may also feel unprepared to cope with acute symptoms. The ever-changing nature and fast-paced workflow in the emergency department (ED) often affect how health care professionals can tailor information and communication to the needs of the patient.

Objective: This study aimed to evaluate the usability and experience of a newly developed information system. The system was developed together with patients and their family members to help provide the information needed in the ED.

Methods: We conducted a mixed methods study consisting of quantitative data obtained from the System Usability Scale questionnaire and qualitative interview data obtained from purposively selected participants included in the quantitative part of the study.

Results: A total of 106 patients and 14 family members (N=120) answered the questionnaire. A total of 10 patients and 3 family members participated in the interviews. Based on the System Usability Scale score, the information system was rated close to excellent, with a mean score of 83.6 (SD 12.8). Most of the participants found the information system easy to use and would like to use it again. The participants reported that the system helped them feel in control, and the information was useful. Simplifications were needed to improve the user experience for the older individuals.

Conclusions: This study demonstrates that the usability of the information system is rated close to excellent. It was perceived to be useful as it enabled understanding and predictability of the patient's trajectory in the ED. Areas for improvement include making the system more usable by older individuals. The study provides an example of how a technological solution can be used to diminish the information gap in an ED context.

(JMIR Hum Factors 2024;11:e48445) doi:10.2196/48445

KEYWORDS

consumer; eHealth; elderly; emergency department; emergency; family members; healthcare professionals; information system; mixed methods research: patients; qualitative interview; questionnaire; technology; usability; usable

XSL•FO
RenderX

Introduction

Background

Clear communication and information are essential to improving care and patient outcomes in the emergency department (ED) [1-5]. A lack of information during ED visits causes patients and their family members to experience a sense of powerlessness and to feel unprepared to cope with acute symptoms [2,3,6]. Due to the hectic nature of the ED and the constant interruptions, communication from health care professionals is often inadequate or not tailored to patients and their families [4,7]. While this problem has been known for many years, it still persists to this date [1].

Health technologies are implemented in many parts of health care systems to promote quality care and treatment [8]. The design and purpose of health technologies range widely from organizational [9] to person-centered intentions [10]. In the ED, technologies may be used as quality dashboards [9] and more personal information systems on patients' own devices to support the delivery of health information [11]. However, the successful use of technology in clinical practice is likely to be ineffective if user needs are not carefully addressed and incorporated before attempting a full-scale implementation [9,12]. Thoroughness in integrating and understanding user perspectives will have a direct impact on how well the technology is suited for clinical practice [13,14].

Based on the current findings, patients in the ED and their family members have unmet information needs [1-4]. Hence, guided by the principles of user-driven activities [15], a health information system was developed [16]. The health information system, which is called "Cetrea Clinical Logistic (CCL) *for patients*," is available for patients in the hospital's emergency room and displays real-time information, including (1) person-centered activities, (2) information videos, (3) a notepad, (4) waiting time, and (5) the nurse and physician responsible for care.

Usability is one of the factors affecting the acceptance of health information systems by users, and it is essential for the effective use of the system [17]. A usability evaluation can identify problems and weaknesses in the design and functionalities in the early development phase [18]. Usability tests allow developers to address and adjust concerns and, thus, avoid implementing technologies that will not be useful in the clinical context.

Therefore, a usability evaluation from an end-user perspective was completed to obtain a nuanced understanding of the sustainable use of the system, specifically from the perspective of patients and their family members.

Objective

The objective of this study is to gain knowledge about the usability and experiences of the newly developed information system, CCL for patients. This study reports on patients' and family members' evaluations of this system.

Participatory Design and Technology

This study is the final phase of a 3-phase participatory design study (Figure 1) [19]. Participatory design is a research methodology based on the epistemological position of genuine involvement and understanding of the needs of future end users. A new technology can be designed to improve a real-life problem [20]. The core principles in participatory design methodology have been the theoretical framework of the overall study. In the initial phase, the author group identified the essential needs of patients in the ED, their family members, and ED clinicians [2,3]. The results from phase 1 informed the second phase, in which an information system, CCL for patients, was developed in a cocreation process [16]. The third phase involved testing and evaluation of the system, which is reported in this study. Reporting the evaluation of participatory-designed health technology is a common part of the research methodology [21,22].

The author group has had no financial interest in the system owners of CCL for patients and has no interest in either marketing or promoting the system.

CCL for patients provides information directly to patients and their family members during their stay in the ED. The information provided relates to treatment and time factors and is adjusted toward the individual patient. CCL is an already existing and implemented system for task management for clinicians' use only [23], whereas CCL for patients is a redesign and further development of the system for patients' use.

The functionalities of the CCL for patients' screen are presented in Figure 2.

XSL•FO
RenderX

Figure 1. Overview of the 3-phase study, highlighting the evaluation phase, which is reported in this study.

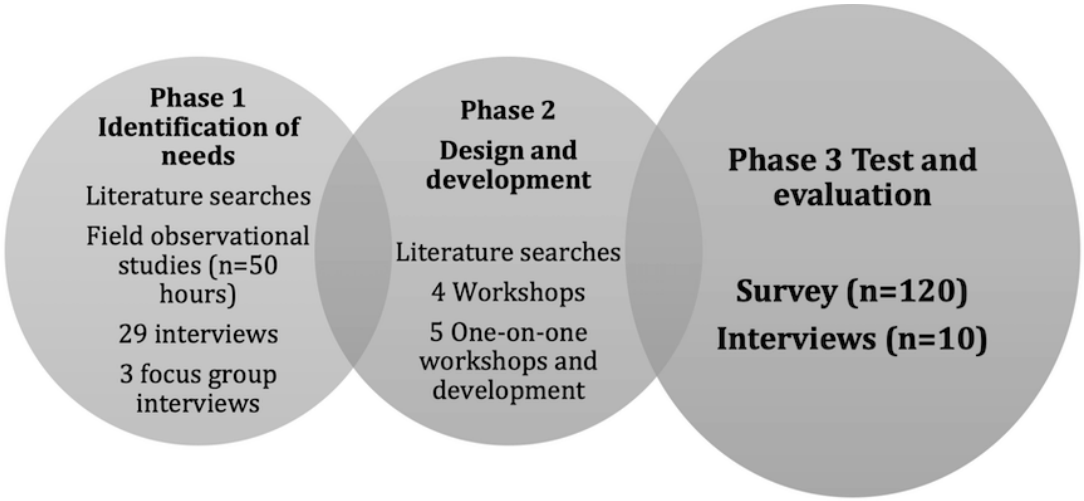
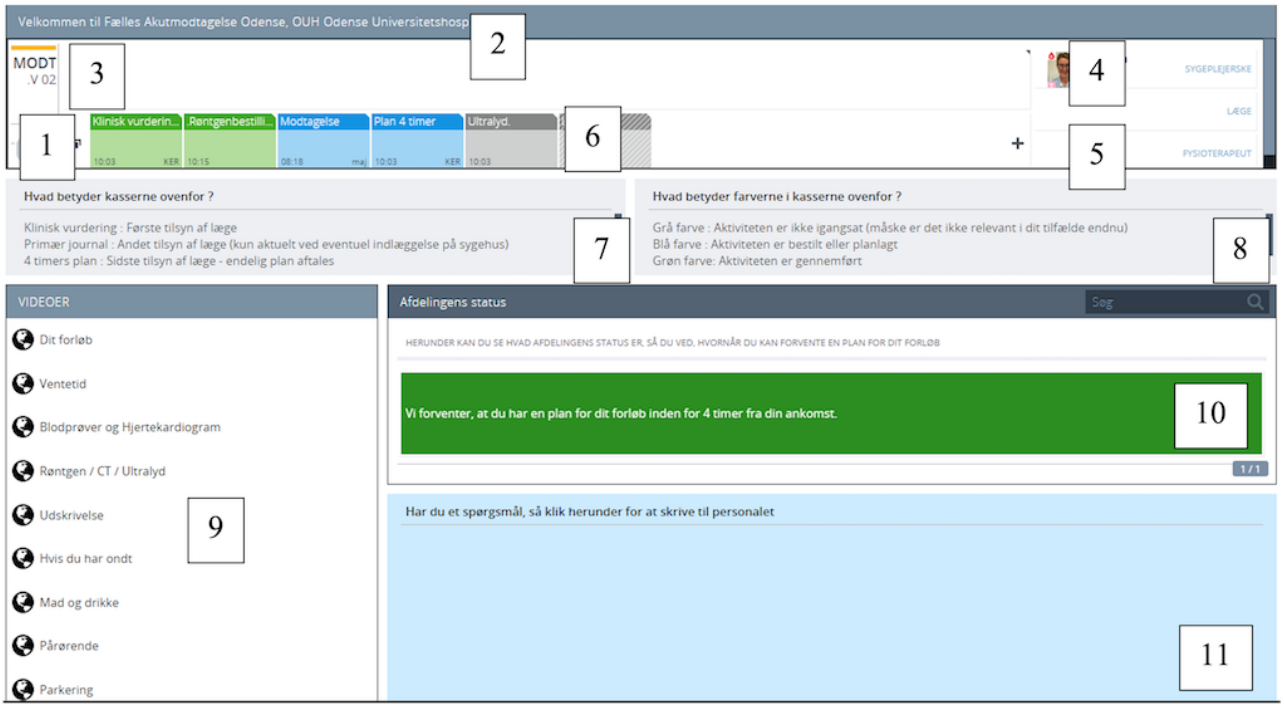


Figure 2. Cetrea Clinical Logistics (CCL) for patients and its functionalities, as displayed to patients and their family members, developed in the second phase of the overall study (the figure has been previously published by Østervang et al [16]). (1) Number of the ED room. (2) Name of the hospital department. (3) The name of the patient (no sensitive information is displayed). (4) The nurse who is responsible for care. (5) The physician who is responsible for treatment. (6) Process line with activities. Displaying nurse assessments, blood samples, electrocardiograms, physician assessments, X-rays, etc. (7) Clarification of the different colors in the process line. Gray: not started; blue: activity scheduled; and green: activity finished. (8) Clarification of special activity names. (9) Link to information videos (eg, information on discharge). (10) Three diverse colors indicate the estimated waiting time: less than 4 hours, equal to 4 hours, or more than 4 hours, respectively. (11) The shared note pad for the patients to write questions to health care professionals or messages from family members.



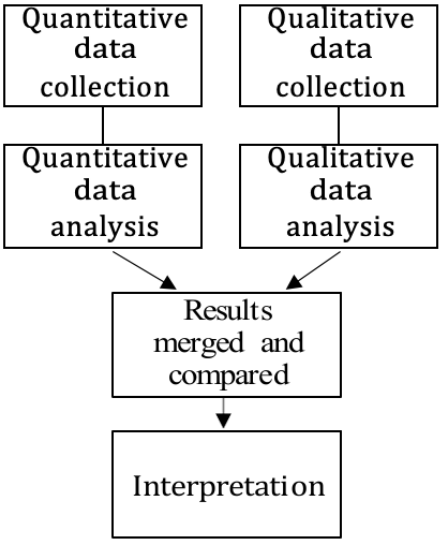
Methods

Research Design

This is a mixed methods study inspired by a convergent parallel design [24]. This design was chosen to obtain nuanced insights into the usability of the system. Further, we adopted this approach to usability testing because quantitative data can identify usability issues and dissatisfaction with program design, while qualitative data can provide detailed information about the causes of the usability issues and point at potential methods for program optimization. As shown in Figure 3 [24], the study contained the following two parts, ending with a merged result: (1) a questionnaire and descriptive characteristics of the participants, and (2) semistructured interviews with patients and their family members.

XSL•FO
RenderX

Figure 3. Diagram for a study using convergent design (Creswell and Clark) [24].



Setting

The data were collected in Odense University Hospital's ED between August 22 and September 29, 2022, on weekdays from 8 AM to 5 PM. The information system was displayed on a laptop personal computer (PC) sitting on the bedside table in the ED room. Four PCs were used during the test phase. They were installed in the specific ED room where the patients participating in the study were admitted.

Inclusion Criteria and Recruitment

All patients admitted to the medical area of the ED without a final plan for treatment and care were eligible for participation. Patients were excluded if they were severely ill or cognitively unable to use the technology. However, patients who were excluded due to a cognitive inability to use the screen but who were still able to give consent for their family members' participation were enrolled if the family member was interested in participating. Patients were recruited by the first author (CØ) or one of 2 research assistants, all of whom have a Master of Nursing Science degree and research experience. Potential participants were identified and discussed with the responsible care nurse before they were approached to reduce the possibility of any concerns.

Quantitative Phase

A survey was conducted to elicit the opinions and experiences of patients and their family members using the information system.

The Questionnaire

The questionnaire, the System Usability Scale (SUS), contained questions regarding the usability of the system. Answers are rated on a 5-point Likert scale from "strongly disagree" to "strongly agree," with 5 representing the highest score (strongly agree) [25]. The participants answered 10 questions from the SUS and 2 questions specific to this study (questions 11 and 12) [25]. These 2 extra questions were added to obtain general information about the participants' experience with CCL for patients (question 11: "I think the system provided a great overview of my stay," and question 12: "I think the information in the system made sense to me"). As SUS has been translated and validated in a Danish hospital context previously (Cronbach α=.87) [26], it was considered suitable for this study.

Sample Size

A total SUS score between 70 and 90 indicates good to excellent usability of the tested system [27]. Based on previous research conducted in Scandinavia using SUS in health care with a reported mean score of 79.81 (SD 14.28), we would gain a 95% CI for a mean score between 77.2 and 82.4 if a total of 120 patients were included [28].

Data Collection

If a patient agreed to participate, the researcher cooperated with the local IT department at the hospital to ensure the patient's access to the system. Initially, the researcher sent the IT department an SMS text message providing information on the PC number and the ED room number. The IT specialist matched the PC and room numbers. Then, the researcher double-checked that the correct information was displayed before handing it to the patient. All participants were given oral guidance on how to use CCL for patients.

The PC with individual information was placed on the bedside table until either the patient left the ED, the patient had used the system for a minimum of 2 hours, or the patient felt ready to perform the evaluation. All of this had to happen no later than 5 PM, when the IT department closed. When returning the PC, the participants were given an iPad to fill out the questionnaire. The data were stored on the logged server OPEN [29], which is part of Odense University Hospital and the University of Southern Denmark.

Qualitative Phase

Interviews were conducted with individual patients or with the patient together with a family member to get a deeper insight into their experiences using the information system.

Interviews

The qualitative part included a subset of the participants from the quantitative part. Before making CCL for patients available

XSL•FO
RenderX

to the participants, they were asked whether they were interested in participating in an interview.

All interviews were conducted by the first author (CØ). By taking a phenomenological-hermeneutical stance, CØ was allowed to recognize her perceptions as an experienced emergency nurse within hermeneutic interpretation [30]. To bridle her preconceived ideas, CØ wrote down her preunderstandings of why patients lack information in the ED. This reflection provided an initial focus for both the overall research question and the interview questions.

The interviews were conducted in the hospital room after the participants had completed the questionnaire. Notes and quotes were taken during the interview. A summary of the conversation was generated at the end of the interview in the form of member checking [31]. A semistructured interview guide inspired by Kvale was used [32]. An example of a question is: "What was your experience of using CCL for patients?" The interviews lasted up to 30 minutes. The interviews were conducted until no new themes arose [33].

Sample Size

To obtain maximal variation, a purposive sampling strategy was used [33]. The inclusion criteria were the same as for the quantitative part of the study, but they also ensured representation of differences in age and gender.

Analysis

Analysis of the Questionnaires

Only fully completed questionnaires were analyzed (N=120). There were no missing data, as the questionnaire was only considered complete if all the questions were answered. According to the SUS guidelines, we performed an individual analysis of each participant's SUS score as well as the mean value for the entire population. We separated the 2 self-constructed questions from the original SUS questions in the calculation and interpretation process to ensure that they were accurate and reliable. The final score was between 0 and 100, where a higher score indicates better usability. Odd-numbered questions were positive in tone, and even-numbered questions were negative in tone, so the scale was converted into points ranging from 1 to 5 (1=strongly disagree to 5=strongly agree). The final score was calculated as follows: X = the sum of the points for all odd-numbered questions minus 5. And Y = 25 minus the sum of the points for all even-numbered questions. SUS score = $(X + Y) \times 2.5$ [34]. A system needs a score above 70 to be considered acceptable; better systems will score from the high 70s to the high 80s, and excellent systems will score above 90 [27].

Analysis of the Interviews

The qualitative interviews were analyzed and reported based on Malterud's [35] systematic text condensation. This process consisted of four steps: (1) transcriptions were read several times to get a total impression of the data and to find preliminary themes; (2) we identified and sorted meaning units based on the preliminary themes and arranged them into code groups; (3) the code groups were reviewed, and the content was reduced into condensates; and (4) the meaning and content of the condensates were synthesized and interpreted [35]. The analysis was completed by CØ using NVivo (version 12; QSR International). The trustworthiness and rigor of the qualitative part of the study were evaluated using Guba's [36] definition of quality criteria. As part of steps 2, 3, and 4 in the analysis, the emerging themes and codes were discussed in the author group toward strengthening the credibility and reflexivity of our interpretation of the interviews. Using a systematic approach toward the analysis strategy of all interviews ensured confirmability in the data collection and analysis process.

The SQUIRE 2.0 checklist [37] was used to create transparency and ensure that no important information was missed in the reporting of the study.

Integration of Quantitative and Qualitative Results

To achieve an expanded understanding of the results, the qualitative and quantitative results were compared and integrated as the final step of the analysis using joint display tables [24]. In a joint display table, the 2 results are presented in a way that allows comparison, leading to confirmation, disconfirmation, or expansion of each other [24]. The results from the SUS (quantitative results) are presented on a Likert scale, showing the variation of the grades in the different questions. To elaborate on and verify the answers, supportive qualitative quotes were presented for each question. We divided the grades into low (1-3) and high (4-5) to separate the different perceptions of CCL for patients.

Ethical Considerations

All the participants received verbal and written information about the study in accordance with applicable ethical rules [38] and provided their oral and written consent. The study is registered with the Danish Data Protection Agency, Fortegnelsen (19/22672). Approval of the project was granted by the Regional Committee on Health Research Ethics for Southern Denmark (S-20192000–111).

Results

Quantitative Results

In total, 14 family members and 106 patients agreed to participate. A total of 27 patients declined to participate for three main reasons: (1) no interest, (2) no technical skills, and (3) a lack of mental ability due to the acute situation.

XSL•FO
RenderX

Table 1. Demographic descriptions of the participants.

Demographic description	Patients (n=106)	Family members (n=14)	Total (N=120)
Gender, n (%)			
Female	55 (51.9)	8 (57.1)	63 (52.5)
Male	51 (48.1)	6 (42.9)	57 (47.5)
Age (years), mean (SD)	55.5 (SD 18.7)	66.5 (SD 11.6)	57 (SD 18.3)
Civil status, n (%)			
No partner	39 (36.8)	2 (14.3)	41 (34.2)
In a relationship	67 (63.2)	12 (85.7)	79 (65.8)
Children, n (%)			
Having children	81 (76.4)	14 (100.0)	95 (79.2)
Having children living at home	32 (39.5)	6 (42.9)	38 (40.0)
Technology, n (%)			
Having a smartphone	96 (90.6)	14 (100.0)	110 (91.7)
Using technology on daily basis	102 (96.2)	13 (92.9)	115 (95.8)
Education, n (%)			
Low	21 (19.8)	1 (7.1)	22 (18.3)
Medium	71 (67.0)	9 (64.3)	80 (66.7)
High	14 (13.2)	4 (28.6)	18 (15.0)

The respondents were equally represented by gender, with a mean age of 57 years. The mean age of family members was higher than that of the included patients. Most participants had medium education levels, but low and high educational levels were also represented.

Overall, the participants answered the survey positively. As displayed in Tables 2 and 3, each item could have a score contribution between 1 and 5. All the odd-numbered (positive) questions had a score contribution above 4.27-4.53, and all the even-numbered (negative) questions had a score ranging from 1.52 to 1.99. Question 1 had the most positive answers: 94.2% (113/120) strongly agreed or agreed that they would like to use the system if they were hospitalized again. Question 4 had the highest negative score value, indicating that the participants felt they needed help using the system. Of the participants, 50.8% (61/120) indicated that they were confident using the system, answering “strongly agree” to question 9, and 87.5% (105/120) strongly agreed or agreed that most people would be able to learn to use this system.

XSL•FO
RenderX

Table 2. Results of the System Usability Scale for all participants (N=120) and the System Usability Scale score contribution of individual items.

System Usability Scale analysis item	Value per 5-point Likert scale response, n (%)					Score contribution (1-5), mean (SD)
	1 (strongly disagree)	2 (disagree)	3 (neutral)	4 (agree)	5 (strongly agree)	
1. I think I would like to use this system, if I am admitted again.	0 (0)	0 (0)	7 (5.8)	42 (35)	71 (59.2)	4.53 (SD 0.61)
2. I found the system unnecessarily complex.	61 (50.8)	41 (34.2)	11 (9.2)	5 (4.2)	2 (1.7)	1.72 (SD 0.92)
3. I thought the system was easy to use.	1 (0.8)	1 (0.8)	8 (6.7)	39 (32.5)	71 (59.2)	4.48 (SD 0.73)
4. I think that I would need help from the staff to be able to use this system.	49 (40.8)	44 (36.7)	11 (9.2)	11 (9.2)	5 (4.2)	1.99 (SD 1.12)
5. I found the various functions in the system to be well correlated.	1 (0.8)	1 (0.8)	8 (6.7)	65 (54.2)	45 (37.5)	4.27 (SD 0.69)
6. I thought there was too much inconsistency in this system.	56 (46.7)	52 (43.3)	8 (6.7)	1 (0.8)	3 (2.5)	1.69 (SD 0.84)
7. I would imagine that most people would learn to use this system very quickly.	0 (0)	0 (0)	15 (12.5)	53 (44.2)	52 (43.3)	4.31 (SD 0.68)
8. I found the system very cumbersome to use.	70 (58.3)	42 (35)	5 (4.2)	2 (1.7)	1 (0.8)	1.52 (SD 0.73)
9. I felt very confident using the system.	2 (1.7)	6 (5)	4 (3.3)	47 (39.2)	61 (50.8)	4.33 (SD 0.89)
10. I needed to learn a lot things before I could get going with this system.	67 (55.8)	42 (35)	7 (5.8)	4 (3.3)	0 (0)	1.57 (SD 0.75)

Based on the answers to the 2 self-constructed questions, (Table 3), 57.5% (69/120) of the participants strongly agreed that CCL for patients provided a great overview of their stay, and 87.5% (105/120) agreed or strongly agreed that the information in the system made sense to them.

Table 3. Results of general questions calculated by System Usability Scale principles for all participants (N=120) and the System Usability Scale score contribution of individual items.

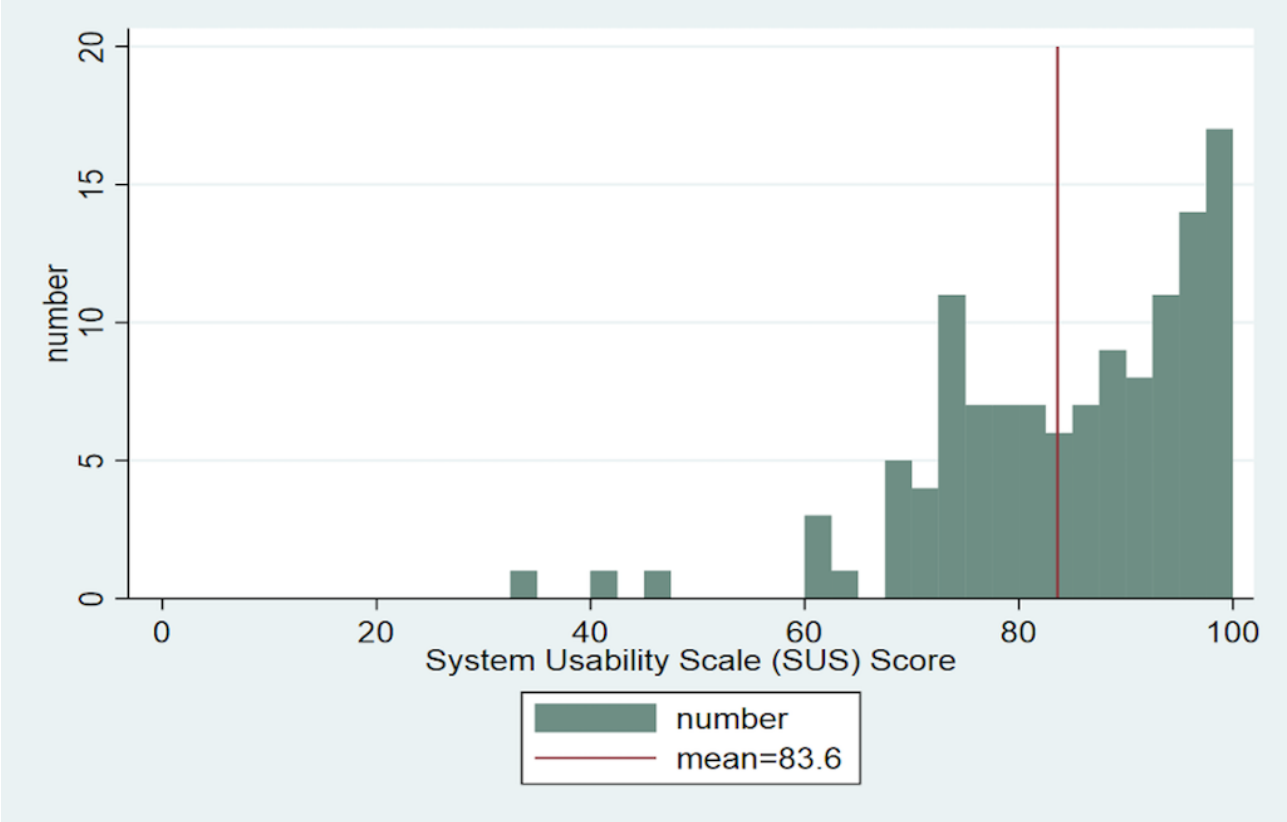
System Usability Scale analysis item	Value per 5-point Likert scale response, n (%)					Score contribution (1–5), mean (SD)
	1 (strongly disagree)	2 (disagree)	3 (neutral)	4 (agree)	5 (strongly agree)	
11. I think the system provided a great overview of my stay.	0 (0)	2 (1.7)	10 (8.3)	39 (32.5)	69 (57.5)	4.46 (SD 0.72)
12. I think the information in the system made sense to me.	1 (0.8)	5 (4.2)	9 (7.5)	45 (37.5)	60 (50)	4.32 (SD 0.85)

For all participants, the total mean score for the SUS scale was 83.6 (SD 12.8), indicating that the system had close to excellent usability.

The median score was 85, and Figure 4 [27] shows the distribution of the individual answers. The scores covered the entire range from 0 to 20 persons per score, and the majority of individuals scored above 70.

XSL•FO
RenderX

Figure 4. Overview of the System Usability Scale (SUS) rating table with inserted value ranges [27].



Qualitative Results

A total of 10 patients and 3 family members (1 daughter aged 55 years, 1 son aged 65 years, and a husband aged 37 years) were interested in elaborating on their experience of CCL for patients after they had tested the system, and the questionnaire was completed. The patients were aged between 32 and 96 years, with equal representation of men and women, and 3 patients were retired.

The following three main themes emerged from the analysis: (1) future perspectives on usability and design; (2) means toward empowerment; and (3) family implications. These themes will be elaborated on using quotes in the upcoming sections.

Future Perspectives on Usability and Design

The majority of the participants expressed a very positive attitude toward CCL for patients but also offered ideas for the future design of the system. The part of CCL for patients that displayed the estimated waiting time in the ED was found to be intuitive and easy to understand and provided informative insights that prepared the participants for their length of stay. This reduced their frustration with not knowing. However, they expressed concerns about the system's lack of familiarity and that it could be improved if the design was like other systems they used in everyday life, such as email or smartphone apps.

> *The system was not difficult at all but I think it would benefit from more recognizability with others systems, for example, email or iPhone applications.* [Male in his 60s]

Most participants valued the line that displayed the boxes with activities the most. They found this part of the system to be essential, as it was the only part that provided direct, personalized information. While they all expressed that they were able to understand the meaning of the changing colors, they also suggested that the text in the boxes could be provided in plain language or a "help" function with text or video could be used to explain the activity in the box.

> *The line with the boxes could be much larger, as this is the most important part! It would be great if you could choose whether you would like to see only the line or all actions on the screen.* [Joint interview, male patient in his 80s and daughter in her 50s]

The participants all watched more than one video, and there was a consensus that the content in the videos was helpful. A few patients who were placed in the hallway due to crowding found it difficult to listen to the videos not only because of the general noise but also because they were afraid of disturbing others. However, the information provided by using videos instead of text was appreciated.

> *The content in the videos was exactly the information I needed. It was nice to be able to revisit the information in the video.* [Female patient in her 50s]

A participant found CCL for patients to be too general. More personalized information, such as individual test results, should be incorporated. Moreover, patients who were visually impaired found the system difficult to use.

> *I have difficulties with my vision, and I do not think I would have been able to use this without help.* [Female patient in her 60s]

XSL•FO
RenderX

Means Toward Empowerment

All of the participants agreed that the system provided an overview that otherwise would not have been accessible for them. Knowing who their treating nurse and doctor were calmed the participants. They described a feeling of not being forgotten in the hectic environment of the ED. Moreover, they valued being able to follow when activities changed from passive to active. Consistency between actions on the screen and in real life provided them with confidence in health care professionals.

> *When you are here, you can hear people working, but you do not know if anyone is taking care of your situation, or you are forgotten. The system helped us to believe we were not forgotten (…) We loved that when something happens on the screen then it was also reflected in real life. E.g. when the screen said the doctor was on his way- he actually came.* [Joint interview, female patient, and husband in their 30s]

Several of the participants stated that having CCL for patients available made them feel calm, as the system provided predictability. Further, having an overview helped them to remain in control of the course of treatment in the ED. Some of the participants said this system could save the nurses' time, as they felt they were more empowered to handle the situation in the ED since they knew what they were waiting for.

> *The questions that I would have needed a nurse to answer were provided by the system; that was really great.* [Female patient in her 30s]

A few of the participants were worried that CCL for patients would need resources from health care professionals that were already scarce.

> *I am worried that the system takes time away from the patients to support the system.* [Joint interview, male patient in his 80s and daughter in her 50s]

Family Implications

Both patients and family members indicated that giving family members direct access was important. CCL for patients gave the family up-to-date information about the care and treatment-related interventions as soon as they attended the hospital room, and they did not need to wait for a nurse or doctor to get an idea of what was planned.

> *My mother is not able to remember what she is planned for today. I think it was great for me to see she is waiting for X rays.* [Female patient in her 90s and son in his 60s]

Furthermore, the family members reported that the system helped them to support the patient, as they could keep track of the interventions provided by CCL for patients. Family members of older patients felt the system was too complicated for the older individuals to use but appreciated that the system was available for them because it allowed them to talk the patient through the stay in the ED. Moreover, a family member stated that the system made it possible for him to let his wife go to sleep, as they agreed that he would wake her up when he saw that activities were about to happen.

> *It was nice for me to have a system that told us when things were going to happen. My wife fell asleep, and I knew I did not need to wake her up before I could see the box turned into the blue color. It was easy to understand.* [Joint interview, female patient, and her husband in their 30s]

Merged Data

We combined the quantitative and qualitative data in a joint display (Multimedia Appendix 1), providing an assessment of the quantitative and qualitative data together. In this way, the data allow us to expand our understanding of patients and their family members' experiences with CCL for patients. For example, in question 1, the participants were asked whether they would like to use CCL for patients again. The participants who gave a lower score (1-3) to that question were concerned if the system would replace personal appearance from health care professionals, whereas those who gave it a high score (4-5) valued how the systems helped them to keep control.

Furthermore, question 7 regarding people's ability to learn to use the system revealed that the participants who gave a low score (1-3) wanted more simplicity, fearing that the older patients would find the system difficult. Meanwhile, the participants who gave high scores (4-5) felt that the system was easy to use. Regarding question 11, the majority of the patients and their family members stated that CCL for patients provided a great overview of the patient's pathway. They further elaborated on this in the interviews, as they felt that the overview of care in the system helped them to feel less stressed and better understand the treatment pathways.

Discussion

Principal Findings

In this study, we report that the perceived overall usability of the health information system CLL *for patients* is good to excellent, providing information that is needed during the entire emergency process. The participants rated the system highly (a score of 83.6 points) and reported that the system gave them an opportunity to remain in control, as they knew what they were waiting for and who was responsible for care and treatment.

Technology as a Means to Empower Patients and Family Members in the ED

Looking into previous research on testing systems using SUS [28], a mean score of 83.6, as found in this study, would indicate that the tested system was successful. However, while CCL for patients was evaluated positively overall, we also uncovered technical concerns regarding usability limitations, specifically regarding the older individuals. Our results showed a mean patient age of 57 years, which represents a relatively young ED population. However, the mean age of the family members was almost 10 years (9.5 years) older. The older individuals found the system to be complicated to use and felt that it needed simplified functions, such as a zoom function and recognizability (eg, other well-known systems). Echoing these findings, Verma et al [39] investigated the level of eHealth literacy among older adults and caregivers and found that one main barrier to the adoption of eHealth was a lack of familiarity with the tools

XSL•FO
RenderX

available. In the development phase of CCL for patients [16], decisions had to be made for the system to work in a clinical setting. One decision was the use of an interface design, which did not allow us to integrate well-known functions, for example, from email or application symbols. Our results highlighted that it might not be possible to design technologies using a one-size-fits-all approach. However, in line with previous research [40], we discovered that the usability testing allowed the developers to adjust and isolate functionalities to provide improved usability outcomes in the future. For example, we found that the participants valued the display with the boxes, which could be promoted in a revised version by the availability of a zoom function.

Furthermore, the participants expressed concerns about whether CCL for patients would influence the health care professionals' available time to provide actual care. Barriers to the adoption of technology systems in clinical settings include the workflow or demand for more human resources [12]. As the information system is a redesigned patient flow system, it would not require changes in workflow or unduly burden professional health care resources. Another consideration was the need for personal test results. They could not be provided in the current form of the system, as it would require a personal log-on to avoid safety issues related to General Data Protection Regulations.

The participants who rated the usability the highest explained that the system made them feel that they were in control of the situation without the fear of being forgotten. The system provided an overview of the care transition and, therefore, offered predictability. This need to be in control has been identified in another study, which described patients' and their relatives' dissatisfaction when visiting the ED [6], as they felt powerless in the ED. Not having knowledge or information available led to such feelings of powerlessness. Nursing rounds were suggested in that study to improve information support [6]. Our results showed that the patients felt more independent because they were able to find the needed information using technology.

Being acutely ill places individuals in a vulnerable situation, and their cognitive capabilities are challenged [2]. Communication from health care professionals and how information is presented have a significant influence on how that information is comprehended [2,41,42]. In this study, we developed information videos related to the journey within the ED, and the participants reported that they were an accessible and usable way to understand information in a stressful situation. Patients and their family members declared that this gave them a feeling of empowerment. Indeed, empowering patients to be in control and involved in their own care is recognized as a core value of high-quality patient-centered care [43]. As Emmamally et al [44-46] noted, improved partnering with family members in the ED is needed. If the family is not included, there is an increased risk of miscommunication and poor understanding of health-related matters [2,44,47-49]. However, creating a closer partnership of care has been described as challenging within the ED due to the high workload, overcrowding, and multitasking [47]. This is echoed in recent findings from studies conducted in a Danish context [2,3], in which family members requested more systematic inclusion in the ED. In this study, the results showed that CCL for patients was perceived as usable and as a useful way to systematically include families during the ED stay.

An update of the Medical Research Council's guidelines for developing and evaluating complex interventions in health care states that appropriate users should be involved in every part of the development, process, and outcome analysis of a complex intervention to ensure sustainable interventions [50]. In line with best practices, the information system has been developed together with representatives of future users of the system, including health care professionals, managers, patients, family members, and IT specialists [16]. For decades, the ED context has been a hectic environment [4,42,51,52]. This creates challenges at both the information and communication levels, affecting whether patients and their families feel in better control during their stay in the ED [1,4,42,51,52]. In this study, we presented and evaluated a simple but unique system that provides timely information to empower individuals without straining health care professionals' resources. The usability test was a crucial and important step to inform changes in functionalities and experiences of using IT in the ED.

Strengths and Limitations

Questionnaires are a common and recognized method for evaluating the usability of health technologies. However, the contextual factors affecting the results are difficult to determine [53]. The SUS did not provide insights on the effectiveness or efficiency of the system, but it is a validated questionnaire and provided an overall understanding of the system [27]. The mixed methods approach [24] enabled the integration of quantitative and qualitative data. This allowed us to obtain an understanding of how the usability was rated and why the results emerged for the specific questions, which is considered a strength of usability testing [40,54].

Additionally, our findings serve as an inspiration to others about how a participatory design process can develop a technology that is aligned with some of the essential needs described by the users of the ED. The findings provide an example of how a technological solution can be used to reduce the information gap in an ED context, as the provision of adequate information to patients and their families is found to be a major challenge in an ED context [2,4,42].

This study also had some limitations. Using a broader evaluation method, for example, a qualitative evaluation questionnaire or an evaluation instrument with more domains, could potentially have provided the study with more nuances [55]. Patients attending the ED outside of the IT department's business hours were not able to use the system. Therefore, we do not know if patients attending the ED in the late evening hours or at night would rate the usability differently. Moreover, no cognitive debriefings or adjustments were made specifically for individuals attending an ED, as these tests were conducted before introducing the questionnaire. Multimedia Appendix 2 [26,56] contains further details about the process as well as final modifications to the questionnaire. In addition, our results are based on a relatively young population (with a mean age of 57 years). Another weakness is that we did not include all users in the evaluation phase, as health care professionals, IT specialists,

XSL•FO
RenderX

and managers were only involved in the development phase and not in the usability testing. For the system to be fully useful, it must run on its own or be serviced directly in the ED. These aspects will be considered in the planning of a future implementation process. Moreover, the transferability of the results is limited to countries with comparable access to and understanding of technologies, as in the Danish population and health care system.

Conclusion

Based on the results of this study, the usability of CCL for patients is rated close to excellent by patients and family members. CCL for patients was perceived to be useful, as it enabled understanding of the ED treatment and pathway. The patients indicated that they, from the technology, were able to understand what was going to happen, experienced the feeling of being in control, and found the information to be useful. Areas for improvement include making the system more usable for the older individuals. It is concluded that a technological solution can be used to minimize the information gap in an ED context from the perspective of patients and their family members.

Acknowledgments

The authors would like to thank all those who participated in this study, OPEN [29], and Niels I Agerbek for data management.

The authors received financial support for the research, authorship, and publication of this study from the Region of Southern Denmark, the University of Southern Denmark, and Odense University Hospital.

Authors' Contributions

CØ, CMJ, KBD, EC, and AL wrote the protocol for the study. CØ and AL performed the data analysis. CØ wrote the first draft of the manuscript. All the authors reviewed and edited the manuscript and approved the final version.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Joint display.
[PDF File (Adobe PDF File), 55 KB - humanfactors_v11i1e48445_app1.pdf]

Multimedia Appendix 2
Supplementary file for the Methods section.
[PDF File (Adobe PDF File), 61 KB - humanfactors_v11i1e48445_app2.pdf]

References

1. Graham B, Endacott R, Smith JE, Latour JM. 'They do not care how much you know until they know how much you care': a qualitative meta-synthesis of patient experience in the emergency department. Emerg Med J 2019;36(6):355-363. [doi: 10.1136/emermed-2018-208156] [Medline: 31003992]
2. Østervang C, Lassen AT, Jensen CM, Coyne E, Dieperink KB. What are the needs and preferences of patients and family members discharged from the emergency department within 24 hours? A qualitative study towards a family-centred approach. BMJ Open 2021;11(11):e050694 [FREE Full text] [doi: 10.1136/bmjopen-2021-050694] [Medline: 34753758]
3. Østervang C, Lassen AT, Øelund K, Coyne E, Dieperink KB, Jensen CM. Health professionals' perspectives of patients' and family members' needs in emergency departments and patient pathway improvement: a qualitative study in Denmark. BMJ Open 2022;12(2):e057937 [FREE Full text] [doi: 10.1136/bmjopen-2021-057937] [Medline: 35228292]
4. Sonis JD, Aaronson EL, Lee RY, Philpotts LL, White BA. Emergency department patient experience: a systematic review of the literature. J Patient Exp 2018;5(2):101-106 [FREE Full text] [doi: 10.1177/2374373517731359] [Medline: 29978025]
5. Hermann RM, Long E, Trotta RL. Improving patients' experiences communicating with nurses and providers in the emergency department. J Emerg Nurs 2019;45(5):523-530. [doi: 10.1016/j.jen.2018.12.001] [Medline: 30655007]
6. Forsgärde ES, Attebring MF, Elmqvist C. Powerlessness: dissatisfied patients' and relatives' experience of their emergency department visit. Int Emerg Nurs 2016;25:32-36. [doi: 10.1016/j.ienj.2015.07.004] [Medline: 26254101]
7. Leventhal EL, Schreyer KE. Information management in the emergency department. Emerg Med Clin North Am 2020;38(3):681-691. [doi: 10.1016/j.emc.2020.03.004] [Medline: 32616287]
8. Currie WL, Seddon JJM. A cross-national analysis of eHealth in the European Union: some policy and research directions. Inf Manag 2014;51(6):783-797. [doi: 10.1016/j.im.2014.04.004]
9. Almasi S, Rabiei R, Moghaddasi H, Vahidi-Asl M. Emergency department quality dashboard; a systematic review of performance indicators, functionalities, and challenges. Arch Acad Emerg Med 2021;9(1):e47 [FREE Full text] [doi: 10.22037/aaem.v9i1.1230] [Medline: 34405145]

XSL•FO
RenderX

10. Du Y, Dennis B, Rhodes SL, Sia M, Ko J, Jiwani R, et al. Technology-assisted self-monitoring of lifestyle behaviors and health indicators in diabetes: qualitative study. JMIR Diabetes 2020;5(3):e21183 [FREE Full text] [doi: 10.2196/21183] [Medline: 32857056]
11. Westphal M, Yom-Tov GB, Parush A, Carmeli N, Shaulov A, Shapira C, et al. A patient-centered information system (myED) for emergency care journeys: design, development, and initial adoption. JMIR Form Res 2020;4(2):e16410 [FREE Full text] [doi: 10.2196/16410] [Medline: 32130144]
12. Østervang C, Vestergaard LV, Dieperink KB, Danbjørg DB. Patient rounds with video-consulted relatives: qualitative study on possibilities and barriers from the perspective of healthcare providers. J Med Internet Res 2019;21(3):e12584 [FREE Full text] [doi: 10.2196/12584] [Medline: 30907746]
13. Kushniruk A, Nøhr C. Participatory design, user involvement and health IT evaluation. Stud Health Technol Inform 2016;222:139-151. [Medline: 27198099]
14. Lau AY, Piper K, Bokor D, Martin P, Lau VS, Coiera E. Challenges during implementation of a patient-facing mobile app for surgical rehabilitation: feasibility study. JMIR Hum Factors 2017;4(4):e31 [FREE Full text] [doi: 10.2196/humanfactors.8096] [Medline: 29217504]
15. Simonsen J, Robertson T. Routledge International Handbook Of Participatory Design. London: Routledge; 2012.
16. Østervang C, Lassen A, Schmidt T, Coyne E, Dieperink KB, Jensen CM. Development of a health information system to promote emergency care pathways: a participatory design study. Digit Health 2022;8:20552076221145856 [FREE Full text] [doi: 10.1177/20552076221145856] [Medline: 36601282]
17. Farzandipour M, Nabovati E, Tadayon H, Jabali MS. Usability evaluation of a nursing information system by applying cognitive walkthrough method. Int J Med Inform 2021;152:104459. [doi: 10.1016/j.ijmedinf.2021.104459] [Medline: 34091145]
18. Bligård LO, Osvalder AL. Enhanced cognitive walkthrough: development of the cognitive walkthrough method to better predict, identify, and present usability problems. Adv Hum Comput Interact 2013;2013:1-17 [FREE Full text] [doi: 10.1155/2013/931698]
19. Østervang C, Lassen AT, Jensen CM, Coyne E, Dieperink KB. How to improve emergency care to adults discharged within 24 hours? Acute care planning in emergency departments (The ACE study): a protocol of a participatory design study. BMJ Open 2020;10(12):e041743 [FREE Full text] [doi: 10.1136/bmjopen-2020-041743] [Medline: 33371037]
20. Kensing F. Methods and Practices in Participatory Design. Copenhagen: ITU Press; 2003.
21. Nielsen C, Agerskov H, Bistrup C, Clemensen J. Evaluation of a telehealth solution developed to improve follow-up after kidney transplantation. J Clin Nurs 2020;29(7-8):1053-1063. [doi: 10.1111/jocn.15178] [Medline: 31889357]
22. Jensen CM, Overgaard S, Wiil UK, Clemensen J. Can tele-health support self-care and empowerment? A qualitative study of hip fracture patients' experiences with testing an "app". SAGE Open Nurs 2019;5:2377960819825752 [FREE Full text] [doi: 10.1177/2377960819825752] [Medline: 33415217]
23. Getinge. 2022. URL: http://www.cetrea.com/index.php/da/ [accessed 2024-01-19]
24. Creswell JW, Clark VLP. Designing and Conducting Mixed Methods Research, Third Edition. Thousand Oaks, California: Sage Publications; 2018.
25. Lewis JR. The system usability scale: past, present, and future. Int J Hum-Comput Int 2018;34(7):577-590. [doi: 10.1080/10447318.2018.1455307]
26. Hvidt JCS, Christensen LF, Sibbersen C, Helweg-Jørgensen S, Hansen JP, Lichtenstein MB. Translation and validation of the System Usability Scale in a Danish mental health setting using digital technologies in treatment interventions. Int J Hum-Comput Int 2019;36(8):709-716. [doi: 10.1080/10447318.2019.1680922]
27. Bangor A, Kortum PT, Miller JT. An empirical evaluation of the System Usability Scale. Int J Hum-Comput Int 2008;24(6):574-594. [doi: 10.1080/10447310802205776]
28. Hägglund M, Scandurra I. User evaluation of the Swedish patient accessible electronic health record: system usability scale. JMIR Hum Factors 2021;8(3):e24927 [FREE Full text] [doi: 10.2196/24927] [Medline: 34313596]
29. OPEN. Open patient data explorative network. Clinical Institute University of Southern Denmark. 2022. URL: https://www.sdu.dk/da/forskning/open [accessed 2024-01-19]
30. Green J, Thorogood N. Qualitative Methods for Health Research, Third Edition. Los Angeles: SAGE Publications; 2014.
31. Birt L, Scott S, Cavers D, Campbell C, Walter F. Member checking: a tool to enhance trustworthiness or merely a nod to validation? Qual Health Res 2016;26(13):1802-1811. [doi: 10.1177/1049732316654870] [Medline: 27340178]
32. Kvale S, Brinkmann S. Interview: introduktion til et håndværk, 2. udgave. Kbh: Hans Reitzel; 2009.
33. Moser A, Korstjens I. Series: practical guidance to qualitative research. part 3: sampling, data collection and analysis. Eur J Gen Pract 2018;24(1):9-18 [FREE Full text] [doi: 10.1080/13814788.2017.1375091] [Medline: 29199486]
34. Brooke J. SUS: a 'quick and dirty' usability scale. In: Jordan PW, McClelland IL, Thomas B, Weerdmeester BA, editors. Usability Evaluation in Industry. Boca Raton, FL: CRC Press; 1996:189-194.
35. Malterud K. Systematic text condensation: a strategy for qualitative analysis. Scand J Public Health 2012;40(8):795-805. [doi: 10.1177/1403494812465030] [Medline: 23221918]
36. Guba EG. Criteria for assessing the trustworthiness of naturalistic inquiries. ECTJ 1981;29(2):75-91. [doi: 10.1007/bf02766777]

XSL•FO
RenderX

37. Ogrinc G, Mooney SE, Estrada C, Foster T, Goldmann D, Hall LW, et al. The SQUIRE (Standards for QUality Improvement Reporting Excellence) guidelines for quality improvement reporting: explanation and elaboration. Qual Saf Health Care 2008;17(Suppl 1):i13-i32 [FREE Full text] [doi: 10.1136/qshc.2008.029058] [Medline: 18836062]
38. World Medical Association. World Medical Association declaration of Helsinki: ethical principles for medical research involving human subjects. JAMA 2013;310(20):2191-2194 [FREE Full text] [doi: 10.1001/jama.2013.281053] [Medline: 24141714]
39. Verma R, Saldanha C, Ellis U, Sattar S, Haase KR. eHealth literacy among older adults living with cancer and their caregivers: a scoping review. J Geriatr Oncol 2022;13(5):555-562. [doi: 10.1016/j.jgo.2021.11.008] [Medline: 34810146]
40. Alwashmi MF, Hawboldt J, Davis E, Fetters MD. The iterative convergent design for mobile health usability testing: mixed methods approach. JMIR Mhealth Uhealth 2019;7(4):e11656 [FREE Full text] [doi: 10.2196/11656] [Medline: 31025951]
41. Samuels-Kalow M, Rhodes K, Uspal J, Reyes Smith A, Hardy E, Mollen C. Unmet needs at the time of emergency department discharge. Acad Emerg Med 2016;23(3):279-287 [FREE Full text] [doi: 10.1111/acem.12877] [Medline: 26683867]
42. Vaillancourt S, Seaton MB, Schull MJ, Cheng AH, Beaton DE, Laupacis A, et al. Patients' perspectives on outcomes of care after discharge from the emergency department: a qualitative study. Ann Emerg Med 2017;70(5):648-658.e2. [doi: 10.1016/j.annemergmed.2017.05.034] [Medline: 28712607]
43. Pekonen A, Eloranta S, Stolt M, Virolainen P, Leino-Kilpi H. Measuring patient empowerment—a systematic review. Patient Educ Couns 2020;103(4):777-787. [doi: 10.1016/j.pec.2019.10.019] [Medline: 31767243]
44. Emmamally W, Brysiewicz P. Families' perceptions of support from health care professionals in the three emergency departments in KwaZulu Natal, South Africa. Int J Africa Nurs Sci 2019;10:55-60 [FREE Full text] [doi: 10.1016/j.ijans.2019.01.004]
45. Emmamally W, Erlingsson C, Brysiewicz P. Partnering with families: Is the ED ready? Int Emerg Nurs 2019;42:1. [doi: 10.1016/j.ienj.2019.01.002] [Medline: 30660319]
46. Emmamally W, Erlingsson C, Brysiewicz P. Families' perceptions of relational practice in the emergency department: a qualitative study. Int Emerg Nurs 2020;51:100877. [doi: 10.1016/j.ienj.2020.100877] [Medline: 32479929]
47. Brysiewicz P, Emmamally W. Focusing on families in the emergency department. Int Emerg Nurs 2017;30:1-2. [doi: 10.1016/j.ienj.2016.08.002] [Medline: 27727139]
48. Hoffmann E, Andersen PT, Mogensen CB, Prinds C, Primdahl J. Relatives' negotiation power in relation to older people's acute hospital admission: a qualitative interview study. Scand J Caring Sci 2022;36(4):1016-1026. [doi: 10.1111/scs.13012] [Medline: 34156115]
49. Hoffmann E, Olsen PR. Like an ace up the sleeve: an interview study of nurses' experiences of the contact with relatives in a somatic emergency ward. Scand J Caring Sci 2018;32(3):1207-1214. [doi: 10.1111/scs.12566] [Medline: 29488643]
50. Skivington K, Matthews L, Simpson SA, Craig P, Baird J, Blazeby JM, et al. A new framework for developing and evaluating complex interventions: update of Medical Research Council guidance. BMJ 2021;374:n2061 [FREE Full text] [doi: 10.1136/bmj.n2061] [Medline: 34593508]
51. Stevens L, Fry M, Browne M, Barnes A. Fast track patients' satisfaction, compliance and confidence with emergency department discharge planning. Australas Emerg Care 2019;22(2):87-91. [doi: 10.1016/j.auec.2019.01.004] [Medline: 31042527]
52. Nairn S, Whotton E, Marshal C, Roberts M, Swann G. The patient experience in emergency departments: a review of the literature. Accid Emerg Nurs 2004;12(3):159-165. [doi: 10.1016/j.aaen.2004.04.001] [Medline: 15234713]
53. Maramba I, Chatterjee A, Newman C. Methods of usability testing in the development of eHealth applications: a scoping review. Int J Med Inform 2019;126:95-104. [doi: 10.1016/j.ijmedinf.2019.03.018] [Medline: 31029270]
54. Rochat J, Ehrler F, Siebert JN, Ricci A, Ruiz VG, Lovis C. Usability testing of a patient-centered mobile health app for supporting and guiding the pediatric emergency department patient journey: mixed methods study. JMIR Pediatr Parent 2022;5(1):e25540 [FREE Full text] [doi: 10.2196/25540] [Medline: 35289754]
55. Kidholm K, Ekeland AG, Jensen LK, Rasmussen J, Pedersen CD, Bowes A, et al. A model for assessment of telemedicine applications: mast. Int J Technol Assess Health Care 2012;28(1):44-51 [FREE Full text] [doi: 10.1017/S0266462311000638] [Medline: 22617736]
56. Wild D, Grove A, Martin M, Eremenco S, McElroy S, Verjee-Lorenz A, et al. Principles of good practice for the translation and cultural adaptation process for patient-reported outcomes (PRO) measures: report of the ISPOR Task Force for Translation and Cultural Adaptation. Value Health 2005;8(2):94-104 [FREE Full text] [doi: 10.1111/j.1524-4733.2005.04054.x] [Medline: 15804318]

Abbreviations

ED: emergency department
PC: personal computer
SUS: System Usability Scale

XSL•FO
RenderX

Edited by P Santana-Mancilla; submitted 26.04.23; peer-reviewed by R Rabiei, R Suomi; comments to author 10.10.23; revised version received 30.11.23; accepted 17.12.23; published 21.02.24.

Please cite as:
Østervang C, Jensen CM, Coyne E, Dieperink KB, Lassen A
Usability and Evaluation of a Health Information System in the Emergency Department: Mixed Methods Study
JMIR Hum Factors 2024;11:e48445
URL: https://humanfactors.jmir.org/2024/1/e48445
doi:10.2196/48445
PMID:38381502

©Christina Østervang, Charlotte Myhre Jensen, Elisabeth Coyne, Karin B Dieperink, Annmarie Lassen. Originally published in JMIR Human Factors (https://humanfactors.jmir.org), 21.02.2024. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on https://humanfactors.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Intention to Use an Electronic Community Health Information System Among Health Extension Workers in Rural Northwest Ethiopia: Cross-Sectional Study Using the Unified Theory of Acceptance and Use of Technology 2 Model

Tesfahun Hailemariam[1], PhD; Asmamaw Atnafu[2], PhD; Lemma Gezie[3], PhD; Jens Kaasbøll[4], Prof Dr; Jorn Klein[5], Prof Dr; Binyam Tilahun[1], PhD

[1]Department of Health Informatics, Institute of Public Health, College of Medicine and Health Sciences, University of Gondar, Gondar, Ethiopia

[2]Department of Health System and Policy, Institute of Public Health, College of Medicine and Health Sciences, University of Gondar, Gondar, Ethiopia

[3]Department of Epidemiology and Biostatistics, Institute of Public Health, College of Medicine and Health Sciences, University of Gondar, Gondar, Ethiopia

[4]Department of Informatics, University of Oslo, Oslo, Norway

[5]Department of Nursing and Health Sciences Campus Porsgrunn, University of South-Eastern Norway, Porsgrunn, Norway

Corresponding Author:
Tesfahun Hailemariam, PhD
Department of Health Informatics, Institute of Public Health
College of Medicine and Health Sciences
University of Gondar
Chechela street
Gondar, 196
Ethiopia
Phone: 251 913305250
Email: tesfahunhailemariam@gmail.com

Abstract

Background: IT has brought remarkable change in bridging the digital gap in resource-constrained regions and advancing the health care system worldwide. Community-based information systems and mobile apps have been extensively developed and deployed to quantify and support health services delivered by community health workers. The success and failure of a digital health information system depends on whether and how it is used. Ethiopia is scaling up its electronic community health information system (eCHIS) to support the work of health extension workers (HEWs). For successful implementation, more evidence was required about the factors that may affect the willingness of HEWs to use the eCHIS.

Objective: This study aimed to assess HEWs' intentions to use the eCHIS for health data management and service provision.

Methods: A cross-sectional study design was conducted among 456 HEWs in 6 pilot districts of the Central Gondar zone, Northwest Ethiopia. A Unified Theory of Acceptance and Use of Technology model was used to investigate HEWs' intention to use the eCHIS. Data were cleaned, entered into Epi-data (version 4.02; EpiData Association), and exported to SPSS (version 26; IBM Corp) for analysis using the AMOS 23 Structural Equation Model. The statistical significance of dependent and independent variables in the model was reported using a 95% CI with a corresponding *P* value of <.05.

Results: A total of 456 HEWs participated in the study, with a response rate of 99%. The mean age of the study participants was 28 (SD 4.8) years. Our study revealed that about 179 (39.3%; 95% CI 34.7%-43.9%) participants intended to use the eCHIS for community health data generation, use, and service provision. Effort expectancy (β=0.256; *P*=.007), self-expectancy (β=0.096; *P*=.04), social influence (β=0.203; *P*=.02), and hedonic motivation (β=0.217; *P*=.03) were significantly associated with HEWs' intention to use the eCHIS.

Conclusions: HEWs need to be computer literate and understand their role with the eCHIS. Ensuring that the system is easy and enjoyable for them to use is important for implementation and effective health data management.

(JMIR Hum Factors 2024;11:e47081) doi:10.2196/47081

XSL•FO
RenderX

KEYWORDS

data capturing; data use; eCHIS; electronic community health information system; health extension worker; HEW; intention to use; service provision; Unified Theory of Acceptance and Use of Technology 2; UTAUT2 model

Introduction

Though IT has demonstrated remarkable promise in closing the digital divide in resource-constrained regions and advancing the health care system, there is a global shortage of health workers, which prevents at least half of the world's population from receiving essential health services [1,2]. Training community health workers (CHWs) in low- and middle-income countries has been a recommended approach to closing the global shortage of health workers [2].

Ethiopia has been implementing the Health Extension Program (HEP) since 2003, comprising female health extension workers (HEWs) to improve the health status of its population [3]. Though various strategies were implemented, and substantial progress was observed in enhancing the community health information system (eCHIS), the performance of HEWs has remained low. The possible reasons for the low performance of HEWs are increased workload; lack of motivation; negligence; and skill gaps in health data production, use, and service provision [4-8].

Due to individual, organizational, and interpersonal level impediments, in most resource-constrained countries, particularly in sub-Saharan regions, health care data generated and used for decision-making are incomplete and inaccurate [9,10]. Likewise, quality health data generation and evidence-based decision-making practices are the remaining challenges for the health care system in Ethiopia [8,11,12].

The growing evidence shows that the penetration of mobile technology improves health service delivery and health outcomes across the world [13-19] and is becoming a solution to strengthen health care industries [8,20-22]. Previous studies in Ethiopia [23], Ghana [24], Uganda [25], South Dakota [26], Indonesia [27], Canada [28], Taiwan [29], South Korea [30], and Jordan [31] indicate that data collection using electronic systems may save time over manual data collection [32,33], and there is the potential to improve health care and the productivity of health staff. For example, digital health solutions may enable CHWs to generate quality health data [34], improve health care delivery [35], and help CHWs be more effective in their job at the community level [32,36].

The eCHIS is one of the evidence-based mobile platforms for CHWs in resource-constrained countries [37], which is an easily customizable mobile health (mHealth) platform for health workers to track and support their interactions with clients. It replaces the conventional practice of a CHW manually tracking their work and carrying large client data and documentation [37].

To tackle the challenges that existed with manual health data generation, use, and service provision, the Ethiopian Federal Ministry of Health has taken the initiative to digitize the existing paper-based Community Health Information System through the eCHIS and started piloting it in 6 districts of the Central Gondar zone, Northwest Ethiopia. The ultimate goal of its implementation is to improve the quality of health data production and service delivery at the community level by transforming the culture of information use by using tablet devices.

The first component of the eCHIS is the HEW component, which supports HEWs in family folder management and the provision of reproductive, maternal, newborn, and child health service delivery and follow-up. The second component is the health center referral component, which enables health center workers to confirm referrals and provide referral feedback to HEWs. The focal person component is the third component, which assists focal persons who are designated at the health center level to provide technical and programmatic support to the HEWs. Therefore, it enables HEWs to manage health post–level data and service provision, as it facilitates referral linkage of clients from health posts to health centers and vice versa.

Although using health system technology has expanded worldwide to leverage quality health data production and use, there is a paucity of evidence on users' behavioral intention to use health system technology [38]. The intention to use a new system is how much a health care provider intends, plans, and predicts their future behavioral readiness to use health care technology [39]. Studies show that users' behavioral intention is one of the significant factors of technology acceptance and use.

Hence, it is critical to evaluate the level of users' intention to use IT before implementing it in the health care system [40-42], as it has a significant role in planning and designing effective implementation strategies for health care programs [43]. Moreover, identifying the level of intention to use the eCHIS for community health data production, use, and service provision and its influencing factors could help to be effective in the implementation and strengthening of the program. To the authors' understanding, the level of HEWs' intention to use the eCHIS for community health care data generation, use, and service provision has not been tested using the Unified Theory of Acceptance and Use of Technology 2 (UTAUT2) model.

The UTAUT2 model is one of the most mature IT models [44] that has emerged from 8 theoretical models that were primarily developed in psychology and sociology [45]. These include the Technology Acceptance Model, Theory of Planned Behavior, Combined Technology Acceptance Model and Theory of Planned Behavior, Theory of Reasoned Action, Motivational Model, Social Cognitive Theory, Model of PC Utilization, and Innovation Diffusion Theory [45,46].

The UTAUT2 has 3 broad types of integration of concepts. First, the integration was examined in new contexts, new users, and new cultural settings [46]. Second, the addition of new constructs increased the scope of dependent predictors [45]. Third, including independent predictors of the Unified Theory

XSL•FO
RenderX

of Acceptance and Use of Technology (UTAUT) variables made comprehension easier [46]. Its extensive replications, applications, and integration extend the theoretical limits of technology adoption. Therefore, the addition of the 3 predictors (hedonic motivation, price value, and habit) to the previously existing 4 constructs in the original UTAUT model (performance expectancy, effort expectancy, social influence, and facilitating conditions) leveraged the adoption and use of technology (eCHIS in this case). This changes the existing relationships of constructs in the original UTAUT and introduces new relationships among constructs known in the UTAUT2.

We used the UTAUT2 constructs to determine HEWs' behavioral intention to use the eCHIS [46], as UTAUT2 perspectives are applicable in the health system and the eCHIS is a form of health system technology. Understanding the intention of HEWs using the UTAUT2 model would give insights to health system leaders on how to digitize community health systems in local settings.

Therefore, this study aimed to investigate HEWs' intention to use the eCHIS and its predicting factors using the UTAUT2 model among HEWs who had received familiarization training on the eCHIS in 6 pilot districts of Northwest Ethiopia.

Since the eCHIS is a form of health system technology, the relationships between UTAUT2 perspectives on accepting and using technology apply to the eCHIS, and the following hypotheses were speculated:

- Hypothesis 1: performance expectancy positively influences HEWs' behavioral intention when using the eCHIS.
- Hypothesis 2: effort expectancy positively influences HEWs' behavioral intention when using the eCHIS.
- Hypothesis 3: social influence positively influences HEWs' behavioral intentions when using the eCHIS.
- Hypothesis 4: facilitating conditions positively influence HEWs' behavioral intentions when using the eCHIS.
- Hypothesis 5: hedonic motivation positively influences HEWs' behavioral intention in using the eCHIS.
- Hypothesis 6: self-efficacy positively influences HEWs' behavioral intention when using the eCHIS.
- Hypothesis 7: habit positively influences HEWs' behavioral intention when using the eCHIS.

In this study, price value was not included in this model because HEWs, the participants in this study, were not directly involved in purchasing the system. Furthermore, the model was not tested on behavioral intention to use the eCHIS among HEWs in Ethiopia.

Methods

Study Design, Period, and Setting

A cross-sectional study design was conducted from January to February 2021 in the Central Gondar zone, Northwest Ethiopia. The Central Gondar zone has 15 districts, of which 6 districts (Wogera, Mirab Dembia, Misrak Dembia, Enfranz, Takusa, and Belesa) were selected as pilot districts for eCHIS implementation in the zone. The estimated total population of the zone was 2,288,440. The zone has a total of 75 health centers and 404 health posts, and there were 897 HEWs (59 urban and 848 rural) during the study period (Central Gondar Zone Health Bureau report, unpublished data, 2020).

Population and Participants of the Study

The source population of the study was HEWs at the primary health care unit level. The study participants were HEWs who were in the pilot districts of the Central Gondar zone and had received initial training for eCHIS implementation. The intervention was skill-oriented training for the implementers of mobile-based community health information system applications based on the training manual prepared by the Ministry of Health, and the training was provided for 1 week by trainers from the regional health bureau and the Ministry of Health. Following the training, each *woreda* (district) led household registration, tablet usage guideline provision, technical support and mentoring, and periodical communications for 1 year.

Provision of Mentorship and Technical Assistance

The University of Gondar assigned three supporting team members who provided technical assistance for implementers with a local mentor every 2 weeks throughout the intervention period. In addition, 1 health information technician (a local mentor) was assigned to provide mentorship and solve eCHIS-related problems during implementation.

Sample Size and Sampling Procedures

The initial sample size was calculated using a single population proportion formula, considering the following assumptions: 50% proportion of intention to use the eCHIS, as there was limited evidence in the area; 95% confidence level for estimations; and 5% margin of error. Using these inputs, the initial sample size was estimated at 385. Considering a 10% nonresponse rate, the final sample size was 422. In the pilot districts, however, the total number of HEWs was 460. Therefore, as the initially determined sample size was closer to the population size, it was planned to include all eligible HEWs in the study.

Study Variables and Measurement

The dependent variable was the intention to use the eCHIS for health data generation and service provision. Based on the UTAUT2, 8 constructs with a 5-point Likert scale were used to assess the intention to use the eCHIS and were considered potential predictors of the study [46].

1. Performance expectancy: the extent to which people believe that using a new technology can improve their job performance [47].
2. Effort expectancy: the degree of ease of use associated with the usage of a new technology [46].
3. Social influence: the degree of importance others recognize in using a new system [45].
4. Facilitating conditions: the degree to which a person perceives that an organization and a technical infrastructure exist to support the intention of people to use technology [45].
5. Hedonic motivation: the motivation to do something due to internal satisfaction [48].

XSL•FO
RenderX

6. Habit: the degree to which users perform the usage of technologies behaviors automatically because of learning [46,49].
7. Self-efficacy: judgment of one's ability to use technology to accomplish a particular job or task [45].
8. Behavioral intention: the degree to which a person has formulated conscious plans to perform or not perform some specified future behaviors [50].

Data Collection Tools and Procedures

Data collection tools were adapted from the source instrument used in the UTAUT2 model [46] in the context of the eCHIS to enhance comprehension by the respondents. The items in the constructs were performance expectancy (4 items), effort expectancy (4 items), social influence (3 items), facilitating condition (4 items), hedonic motivation (3 items), self-expectancy (4 items), habit (4 items), and intention to use (3 items). The source language of the instrument was translated forward into the local language of Amharic, and a backward translation was done to ensure the consistency of the tool. Experts with health management information system backgrounds were invited to review the relevance of each question in the instrument. The experts reviewed the instrument and checked its content and face validity, and the instrument was refined according to the comments given. A pretest was conducted on 5% of the study participants before actual data collection was started, and the tool was refined based on the pretest results. A total of 4 data collectors and 2 supervisors were recruited and trained on the purpose, tools, and procedures of the study. Self-administrated questionnaires were used to collect data from HEWs with the assistance of data collectors and supervisors. The data collection period was from January 28 to February 13, 2021, after 2 weeks of eCHIS familiarization training had been given.

Data Management and Analysis

The data were entered into Epi-data (version 4.02; EpiData Association) and exported to SPSS (version 26; IBM Corp) for descriptive statistics such as frequency, cross-tabulations, and univariate analysis of sociodemographic and model constructs. Simple and multiple structural equation models were carried out using the AMOS 23 Structural Equation Model in order to test the relationship between observed and latent variables and identify the predicting variables of the intention to use the eCHIS. During analysis, we applied a parceling technique to increase model efficiency [51]. The subset-item-parcel approach was used in order to aggregate items into several parcels and use them as indicators of the target construct [52]. Accordingly, we created 2 parcels for each factor of target latent constructs (such as performance expectancy, effort expectancy, facilitator conditions, self-expectancy, and habit) by aggregating randomly grouped items within each scale [53]. The remaining 3 latent target constructs with 3 indicators per construct, such as social influence, hedonic motivation, and intention to use, remained as they existed in the original UTAUT2 model [46].

The overall model's fitness was measured and assessed using the goodness of fit indices such as chi-square ratio (<3), the goodness of fit index (>0.9), adjusted goodness of fit index (>0.8), normal fit index (>0.9), comparative fit index (>0.9), Tucker-Lewis index (>0.9), and root mean square error of approximation (<0.08). For the structural equation model, standardized path coefficients of the regression weight values were used to estimate the path coefficients of the dependent and independent variables. Standardized coefficients are not dependent on the scales as they vary from –1 to 1, where 0 indicates no relationship, 1 indicates a strong positive relationship, and –1 indicates a strong negative relationship. A critical ratio (regression weight or standard error) was used to evaluate whether the constructs had a significant relationship. The absolute value of a critical ratio greater than 1.96 is an indication of the significance of the path coefficients. In this study, the CI and its *P* value were calculated using bootstrapping, and the statistical significance of dependent and independent variables in the model was reported using a 95% CI with a corresponding *P* value of <.05 (Table 1). The square multiple correlation (R^2) was used to report the proportion of variance so that the intention to use the eCHIS could be explained by the model.

Table 1. Structural equation modeling fitness for intention to use electronic community health information system among health extension workers in Northwest Ethiopia, 2021.

Fit indices	Threshold value	Authors	Results obtained	Conclusion
Chi-square	<3	Bentler [54] (1990)	2.67	Accepted
Goodness-of-fit index	>0.9	Chau [55] (1997)	0.92	Accepted
Adjusted goodness-of-fit index	>0.8	Chau [55] (1997)	0.88	Accepted
Comparative fit index	>0.9	Bentler [54] (1990)	0.97	Accepted
Root mean square error of approximation	<0.05	Browne and Cudeck [56] (1993)	0.08	Accepted
Normed fit index	>0.9	Bentler and Bonett [57] (1980)	0.95	Accepted

Reliability and Validity of the Research

Regarding the reliability and validity of the study, Cronbach α reliability coefficients were computed to determine the internal consistency of the constructs. Cronbach α of .7 or above indicates high reliability; between .5 and .7 indicates moderate reliability; and less than .5 indicates low reliability. We have used 4-item Likert questions to assess the reliability of the constructs. Accordingly, the reliability of the constructs assessed by 3-item questions as follows: performance expectancy (α=.92), effort expectancy (α=.87), facilitating condition (α=.75), self-expectancy (α=.88), habit (α=.84), social influence

XSL•FO
RenderX

(α=.78), hedonic motivation (α=.90), and intention to use eCHIS (Table 2). In this study, the magnitude of intention to use the eCHIS was assessed by a 3-item Likert question with a reliability test of Cronbach α=.93.

Table 2. Reliability of the constructs on intention to use the electronic community health information system (eCHIS) among health extension workers in Northwest Ethiopia, 2021.

Constructs	Sample size	Number of items	Cronbach α
Performance expectancy	456	4	.92
Effort expectancy	456	4	.87
Social influence	456	3	.78
Facilitating conditions	456	4	.75
Hedonic motivation	456	3	.90
Self-expectancy	456	4	.88
Habit	456	4	.84
Intention to use the eCHIS system	456	3	.93

Ethical Considerations

Study approval and ethical clearance were obtained from the University of Gondar's ethical review board (R.NO. V/P/RCS/05/2020) and a support letter from the ethical review committee of the Amhara Regional Health Bureau Research and Technology transfer office. Study permission was sought at all levels of governmental administration systems including health offices and health facilities. Written consent was obtained, and participants were informed about the objective, importance of the study, procedure and duration, risk and discomfort, benefits of participating in the study, confidentiality, and the right to refuse or withdraw during data collection. To ensure confidentiality, their names and other personal identifiers were not registered. Participants were not compensated for study participation. We confirm that the provided ethics approval documentation covers the study presented in this manuscript.

Results

Sociodemographic and Other Characteristics of the Study Participants

A total of 456 HEWs participated in the study, with a response rate of 99%. The mean age of the study participants was 28 (SD 4.8) years. More than two-thirds (n=314, 68.9%) of the study participants had work experience of more than 5 years. About half of the participants (n=232, 50.9%) were level 4 (10+4) in their educational status, and the majority of the respondents (n=307, 67.3%) were married. The number of HEWs who had difficulties recharging mobile phones was 307 (67.3%). Our study found that 147 (32.2%) HEWs used Microsoft applications daily , 331 (72.6%) had experience using mobile phones for more than 5 years and above, and 421 (92.3%) had informal mobile phone usage practices or were using personal mobile for health post–related activities (Table 3).

According to the findings of this study, 122 (26.8%), 132 (28.9%), and 162 (35.5%) HEWs strongly agreed to intend, predict, and plan to use the eCHIS, respectively (Table 4).

XSL•FO
RenderX

Table 3. Sociodemographic and informal phone use characteristics of the study participants in Northwest Ethiopia, 2021.

Variables and categories	Values (N=456), n (%)
Age group (years)	
<24	104 (22.8)
25-34	295 (64.7)
≥35	57 (12.5)
Marital status	
Married	307 (67.3)
Single	118 (25.9)
Divorced	26 (5.7)
Widow	5 (1.1)
Work experience (years)	
0-2	55 (12.1)
3-5	87 (19.1)
>5	314 (68.9)
Level of education	
Level I	5 (1.1)
Level II	12 (2.6)
Level III	196 (43)
Level IV	232 (50.9)
Others[a]	11 (2.4)
Difficulty with battery recharging	
Yes	307 (67.3)
No	149 (32.7)
Using Microsoft applications for work and daily life	
Yes	147 (32.2)
No	309 (67.8)
Do you use personal mobile phone for health post–related activities?	
Yes	421 (92.3)
No	35 (7.7)
For how long you have used mobile phone (years)?	
0-5	125 (27.4)
>5	331 (72.6)

[a]Health extension worker with additional diploma, BSc degree, or both.

Table 4. Health extension workers' intention to use the electronic community health information system (eCHIS) in Northwest Ethiopia, 2021 (N=456).

Items	Strongly disagree, n (%)	Disagree, n (%)	Neutral, n (%)	Agree, n (%)	Strongly agree, n (%)
I intend to use the eCHIS system in the future	7 (1.5)	12 (2.6)	12 (2.6)	303 (66.4)	122 (26.8)
I predict I will use the eCHIS system in the future	5 (1.1)	13 (2.9)	9 (2)	297 (65.1)	132 (28.9)
I plan to use the eCHIS system in the future	4 (0.9)	11 (2.4)	15 (3.3)	264 (57.9)	162 (35.5)

XSL•FO
RenderX

Mean Score of All Predictors and Intention to Use the eCHIS Using the UTAUT2 Model

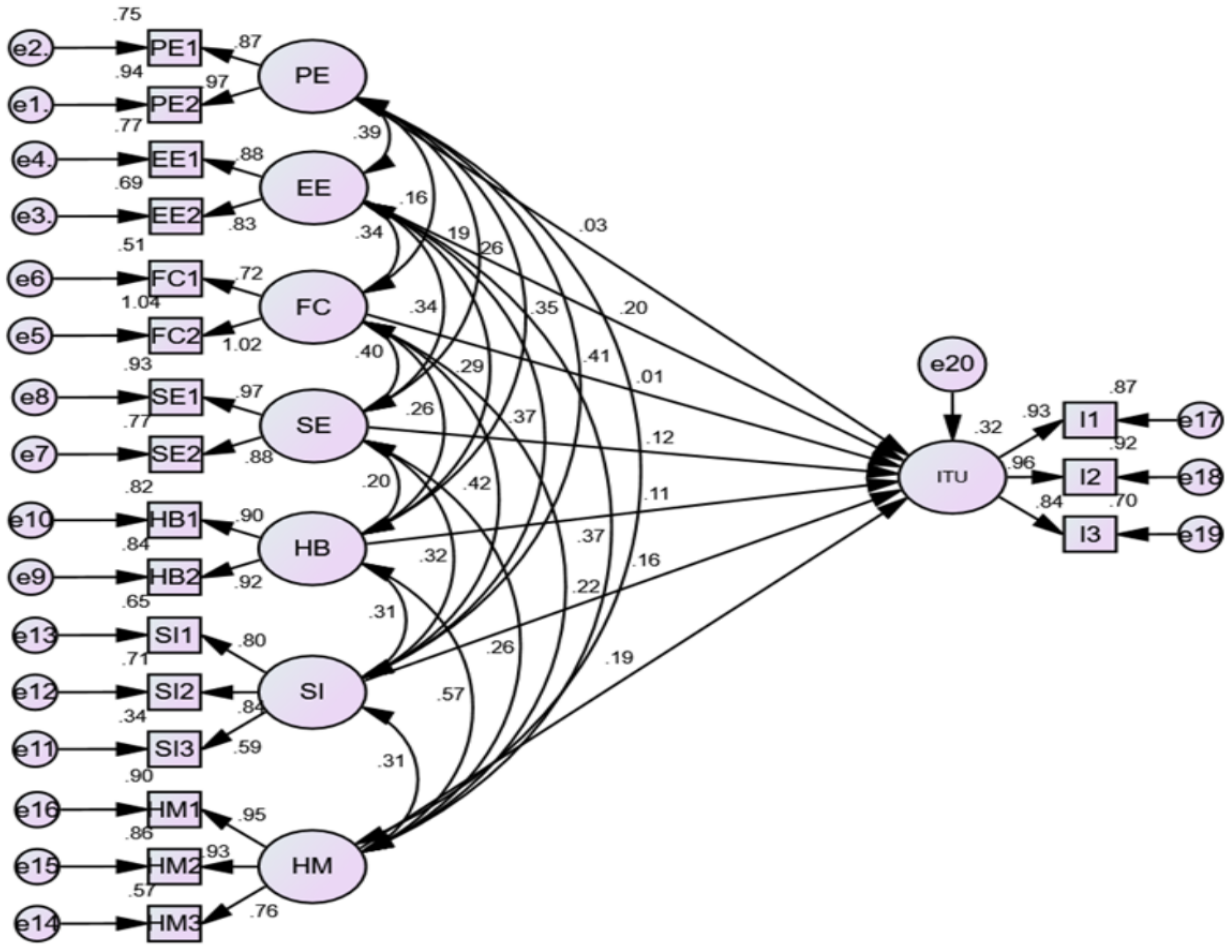
The mean scores of performance expectancy, effort expectancy, facilitating condition, self-expectancy, and habit with 4-item Likert questions were 17.09 (SD 2.58), 16.22 (SD 2.41), 12.07 (SD 3.46), 13.94 (SD 3.62), and 14.75 (SD 3.14), respectively. On the other hand, social influence, hedonic motivation, and intention to use the eCHIS with 3-item Likert questions had a mean score of 11.63 (SD 2.32), 12.23 (SD 2.01), and 12.57 (SD 2.00), respectively.

Our study revealed that 179 (39.3%; 95% CI 34.7%-43.9%) participants who had the intention to use the eCHIS for community health data generation, use, and service provision had scored above the mean. The mean score of the intention to use the eCHIS was 12.57 (SD 2.00). The maximum score of intention to use the eCHIS was 15, while the minimum score was 3.

Simple Structural Equation Model Analysis

The variance in the dependent variable explained by the independent variables was interpreted using square multiple correlation (R^2). The overall R^2 of the intention to use the eCHIS is found to be 32%, the variance that was explained by the independent variables in the model. The bootstrap method with a 95% bias-corrected CI was applied to investigate the significance of path coefficients and factors predicting the model. The predictors with $P<.20$ in the simple structural equation model were considered candidate variables for multiple structural equation model analysis. Due to its undependability to scale, we used a standardized beta coefficient to interpret the influence of predictors on the intention to use the eCHIS. A 95% CI with $P<.05$ was considered to declare an association between dependent and independent variables. The study indicated that effort expectancy has the highest direct effect on HEWs' intention to use the eCHIS, followed by hedonic motivation. The remaining model constructs that have a direct influence on predicting intention to use the eCHIS are social influence and self-expectancy. The structural equation model predicted, with the path coefficients and R^2, is represented in Figure 1, and the path coefficients and P value found from the depicted model are presented in the *Results* section. Moreover, the absolute value of the critical ratio of effort expectancy (3.701), self-expectancy (2.468), social influence (2.782), and hedonic motivation (3.311) indicated that predictors had a significant influence on HEWs' intention to use the eCHIS. Overall, 32% of the variance with respect to intention to use the eCHIS was reasonably explained by the predictors in the model.

Figure 1. Predictors and intention to use the electronic community health information system among health extension workers at Central Gonda zone, Northwest Ethiopia, 2021. EE: effort expectancy; FC: facilitating condition; HB: Habit; HM: hedonic motivation; ITU: intention to use; PE: performance expectance; SE: self-expectancy; SI: social influence.



XSL•FO
RenderX

Effort expectancy (the extent to which people believe that using the eCHIS can improve their effort) has a positive influence on HEWs' behavioral intention (β=.256; *P*=.007). Similarly, self-efficacy and social influence had a positive influence on HEWs' behavioral intention (β=.096; *P*=.04), and (β=.203; *P*=.02), respectively. Likewise, hedonic motivation to use eCHIS due to internal satisfaction was found to be (β=0.217; *P*=.03) and had a significant effect on intention to use the eCHIS. Facilitating conditions (β=0.005; *P*=.92), habit (β=0.103; *P*=.07), and performance expectancy (β=0.034; *P*=.61) had no significant influence on intention to use the eCHIS (Table 5).

Table 5. Multiple structural equations modeling association between predictors and intention to use the electronic community health information system among health extension workers in Northwest Ethiopia, 2021.

Hypothesis	Estimate	95% CI	*P* value	Decision
PE[a] ⇒ IU[b]	0.034	–0.91 to 0.190	.61	Not supported
EE[c] ⇒ IU	0.256	0.060 to 0.503	.007	Supported
FC[d] ⇒ IU	0.005	–0.060 to 0.079	.92	Not supported
SE[e] ⇒ IU	0.096	0.004 to 0.191	.04	Supported
HB[f] ⇒ IU	0.103	–0.008 to 0.235	.07	Not supported
SI[g] ⇒ IU	0.203	0.039 to 0.390	.02	Supported
HM[h] ⇒ IU	0.217	0.022 to 0.416	.03	Supported

[a]PE: performance expectance.
[b]IU: intention to use.
[c]EE: effort expectancy.
[d]FC: facilitating condition.
[e]SE: self-expectancy.
[f]HB: habit.
[g]SI: social influence.
[h]HM: hedonic motivation.

Discussion

Principal Findings

In this study, nearly 2 out of 5 HEWs had an intention to use the eCHIS for community health data generation, use, and service provision. Effort expectancy, self-expectancy, social influence, and hedonic motivation were statistically significant predictors of intention to use the eCHIS. The intention to use the eCHIS by HEWs could be associated with the fact that using the eCHIS is not difficult to understand. It saves time and reduces the amount of effort required to complete health-related tasks [58,59]. Furthermore, it simplifies activities and helps them access data easily. The other could be people around them who have the ability to influence their intention to use the system [58]. For example, HEWs' activities should be monitored and evaluated by health system leaders. If they give them more attention, they will be encouraged to use the system. The other could be previous exposure to using informal phone for health system activities, such as reminding clients about their health care appointments and facilitating referral linkage between health centers and health posts, as mHealth enhances communication between health workers and clients [58]. Moreover, using the eCHIS creates a conducive environment for HEWs since their usual data handling approach is exhaustive and takes much time to execute activities at the health post level, and using the eCHIS not only helps them to save their time but also creates motivation to do their job at health post level [58,60].

Regarding factors associated with intention to use the eCHIS, effort expectancy had a positive influence on the intention to use the eCHIS among HEWs. This finding was in accordance with a study conducted in Ethiopia [61], Kenya [61], the United States [62], and Portugal [63] and had a positively significant association with the intention of health care providers to use technology. A possible explanation could be the fact that the less effort the user devotes to using the system, the more likely he or she is to continue to use it. A study in this regard showed that individuals often want to face a system that is easy to use [64]. HEWs might perceive that the eCHIS could help them to do their job aids shortly with less strain and increased work efficiency [65], as using the eCHIS would simplify the tasks they are expected to deliver at health post level. A review in this regard showed that using digital tools simplifies work and helps to access data easily [59]. Furthermore, studies indicate that digital health solutions reduce workload and improve work performance [24,25], reduce errors [34], create motivation and learning opportunities [66], promote health care appointment [67], and are easy to use and improve work efficiency [59]. Using the eCHIS could reduce the workload of HEWs since manual data management practice at health post level is exhaustive and takes much time to collect data and conduct routine activities [59]. Moreover, the referral linkage integrated into the eCHIS, including HEWs, midwives, and focal persons, will harmonize HEWs' activity flow from health posts to health centers and vice versa. Furthermore, using mHealth motivates

XSL•FO
RenderX

CHWs and enables them to perform multiple tasks quickly, reducing efforts and improving performance [60].

The intention to use the eCHIS among HEWs who perceived people around them could influence their behavioral intention was positively associated. The current finding corroborates studies conducted among health care providers using the UTAUT2 model in Ethiopia [61], Morocco [68], Taiwan [62], South Korea [30], and the United States [69], showing that social influence significantly predicted health care providers' intention to use technology. The possible explanation could be that HEWs might perceive peer pressure from health care staff at the *woreda* and facility levels toward using the eCHIS, which could positively influence their intention to use the eCHIS. The other justification could be the fact that HEWs might get trust from the community for the job aids or activities they are expected to deliver. Hence, health system staff need to understand that peer influence has a positive effect on using a new system. Moreover, making people aware of a new system at the *woreda* and facility levels in general and at the kebele leaders, women's development army, and voluntary service providers' levels, in particular, could influence HEWs' behavioral intention to use the eCHIS. A study in this regard showed that the more health workers connected to colleagues, the more they improved the use of digital tools and the quality of care [58].

Our study revealed that the magnitude of intention to use the eCHIS among HEWs who had self-expectancy was positively correlated. The findings of past studies in Ethiopia [61], Malaysia [70], Taiwan [71], and Iran [72] showed that digital literacy was correlated with the intention to use technology in health care industries. The possible reason might be that those who had self-expectancy could not face difficulty in adapting the emerging technology to community-level data management and service provision. The current evidence in the feasibility and effectiveness study on digital health indicated that the level of computer literacy had influenced digital health implementation among CHWs [73]. A possible explanation might be the fact that informal mobile phone usage practices of CHWs for health post–related activities could influence behavioral intention to use the eCHIS. A study indicated that in many different settings, CHWs use their personal phone informally for community-based activities so as to fill the gaps in the health care system [74].

Our study revealed that there is a significant association between intention to use the eCHIS and hedonic motivation or perceived enjoyment from using the eCHIS for community health data generation, use, and service provision. A possible explanation could be the fact that using a new system instead of the usual approach to manage community-level data and service provision may create intrinsic motivation for HEWs to obtain fun or pleasure. A study showed that motivation is an important construct for eHealth users, and it could even be a sufficient reason to adopt newly emerging technology in a contextual environment [75]. In addition, using eHealth technology to deal with community health data generation, use, and service provision may be an enjoyable process and will have a positive influence on the behavioral intentions of the users [72].

HEWs were optimistic about using the eCHIS because it could be related to the production of quality health data, ease of data management, reduced errors and false reports, data protection, and increased accessibility. A study also indicated that using digital tools could enhance the productivity of CHWs [76]. Community health digitization using mobile apps support the services delivered by CHWs [77]. Furthermore, studies show that the digitization of health care data has promising results in improving both health care and health outcomes [13-19] and improving health staff productivity and work efficiency [65]. In our study, the level of users' optimistic perceptions of using the eCHIS could be an advantage in implementing the intervention, as compared to the existing approach [78]. Studies showed that digital health solutions enable CHWs to generate quality health data [34], improve health care delivery [35], and help them to be more effective in their job aids [32,36]. Moreover, digital tools could help them follow the correct order or the required service elements that clients should receive when providing services. It also enables them to communicate with clients in a better way as compared to manual communication since the tool has prespecified data elements that should be asked by CHWs during service delivery [58].

Likewise, it creates enjoyment among users [66] and benefits them by keeping data safe from human and natural factors that could damage the data. In addition, enjoyment could emanate due to the fact that using digital tools can improve data capturing, storing, and reporting of more items that could be more time-consuming during manual data handling and reduce the motivation of health workers to keep data recording [59]. Even though HEWs are optimistic to use the eCHIS, lack of adequate resources for eCHIS implementation at the implementation district could hinder its successful implementation, and therefore resource availability is vital to be effective in community health digitization. Studies show that challenges during digital health solutions implementation, such as the initial and ongoing capacity-building training [73], poor network access and poor access to electricity [58,79], low financial investment [73], and unreliability or absence of infrastructure (eg, electricity and network) [80,81], hinder the implementation. As the skills and knowledge of HEWs vary from one to another, there should be mentoring and supportive supervision during the implementation. Studies showed that the inability to use the system could affect its implementation [26], and intensive training with continuous refreshment could help them realize the digitization of the community health information system [82].

Limitations of the Study

The findings of this study should be interpreted in light of some limitations. Due to the nature of the study, which was cross-sectional, the inability to infer cause-and-effect relationships is present. As the study was focused on HEWs' intention to use the eCHIS in a pilot district in Northwest Ethiopia, the sample size could affect the findings of this study, and covering larger areas at the regional and national levels is possible, including urban HEWs. Finally, the parcel approach used in this study may introduce parameter estimation bias.

XSL•FO
RenderX

Conclusion

In conclusion, 39.3% (179/456) of HEWs scored above the mean of intention to use the eCHIS for community health data management and service provision. Factors associated with the intention to use the eCHIS were effort expectancy, self-expectancy, social influence, and hedonic motivation. The eCHIS has numerous advantages and a promising future in terms of improving data quality, use, and service delivery. Its adoption in the country, however, should focus on identifying all necessary prerequisites for successful implementation and advancing the community health information system. The implementation of the eCHIS should not skip factors that had no significant effect on intention to use the eCHIS, and further studies at the regional and national levels are recommended to investigate their correlation with intention to use the eCHIS. Model explainability was found in the study using factors that existed in the UTAUT2 model; however, it is recommended to examine the moderating effects of CHWs' related variables to examine how the model constructs could influence HEWs' intention to use the eCHIS.

Acknowledgments

We would like to thank the Central Gondar Zone Health Department, electronic community health information system pilot district health offices, and the University of Gondar Comprehensive Specialized Hospital for their provision of necessary information and support during data collection. Our gratitude also goes to the study participants, data collectors, and supervisors who took part in the study. This work would not have been possible without the financial support of the Doris Duke Charitable Foundation (grant 2017187). The mission of the Doris Duke Charitable Foundation is to improve the quality of people's lives through grants supporting the performing arts, environmental conservation, medical research, and child well-being, and through the preservation of the cultural and environmental legacy of Doris Duke's properties.

Data Availability

The data sets generated and analyzed during this study are available from the corresponding author on reasonable request.

Authors' Contributions

TH conducted the study. All authors conceptualized the design of the study, provided a review of the methodology and results, contributed to the reviewing the manuscript, and read and approved the final manuscript.

Conflicts of Interest

None declared.

References

1. Tamrat T, Kachnowski S. Special delivery: an analysis of mHealth in maternal and newborn health programs and their outcomes around the world. Matern Child Health J 2012;16(5):1092-1101 [FREE Full text] [doi: 10.1007/s10995-011-0836-3] [Medline: 21688111]
2. WHO guidelines on health policy and system support to optimize community health worker programmes. World Health Organization. 2018. URL: https://www.who.int/publications/i/item/9789241550369 [accessed 2024-02-15]
3. Haregu TN, Alemayehu YK, Alemu YA, Medhin G, Woldegiorgis MA, Fentaye FW, et al. Disparities in the implementation of the health extension program in Ethiopia: doing more and better towards universal health coverage. Dialogues Health 2022;1:100047 [FREE Full text] [doi: 10.1016/j.dialog.2022.100047] [Medline: 22489612]
4. Teklehaimanot HD, Teklehaimanot A. Human resource development for a community-based health extension program: a case study from Ethiopia. Hum Resour Health 2013;11(1):39 [FREE Full text] [doi: 10.1186/1478-4491-11-39] [Medline: 23961920]
5. Medhanyie A, Spigt M, Dinant G, Blanco R. Knowledge and performance of the Ethiopian health extension workers on antenatal and delivery care: a cross-sectional study. Hum Resour Health 2012;10(1):44 [FREE Full text] [doi: 10.1186/1478-4491-10-44] [Medline: 23171076]
6. Koblinsky M, Tain F, Gaym A, Karim A, Carnell M, Tesfaye S. Responding to the maternal health care challenge: the Ethiopian health extension program. Ethiop J Health Dev 2010;24(1):105-109 [FREE Full text] [doi: 10.4314/ejhd.v24i1.62951]
7. Teklehaimanot A, Kitaw Y, G/Yohannes A, Girma S, Seyoum S, Ye-Ebiyo Y. Study of the working conditions of health extension workers in Ethiopia. Ethiop J Health Dev 2007;21(3):246-259 [FREE Full text] [doi: 10.4314/ejhd.v21i3.10055]
8. Sebastian MS, Lemma H. Efficiency of the health extension programme in Tigray, Ethiopia: a data envelopment analysis. BMC Int Health Hum Rights 2010;10(1):16 [FREE Full text] [doi: 10.1186/1472-698X-10-16] [Medline: 20546597]
9. Theo L, Rainer S, Bodart C. A framework for designing health information systems. World Health Organization. 2000. URL: https://www.afro.who.int/sites/default/files/2017-06/AHO_Country_H_Infos_Systems_2nd_edition.pdf [accessed 2024-02-15]

XSL•FO
RenderX

10. Tilahun B, Derseh L, Atinafu A, Mamuye A, Mariam TH, Mohammed M, et al. Level and contributing factors of health data quality and information use in two districts in Northwest Ethiopia: social-ecological perspective. BMC Med Inform Decis Mak 2021;21(1):373 [FREE Full text] [doi: 10.1186/s12911-021-01741-1] [Medline: 34972511]
11. Medhanyie A, Spigt M, Kifle Y, Schaay N, Sanders D, Blanco R, et al. The role of health extension workers in improving utilization of maternal health services in rural areas in Ethiopia: a cross sectional study. BMC Health Serv Res 2012;12(1):352 [FREE Full text] [doi: 10.1186/1472-6963-12-352] [Medline: 23043288]
12. Dynes M, Buffington ST, Carpenter M, Handley A, Kelley M, Tadesse L, et al. Strengthening maternal and newborn health in rural Ethiopia: early results from frontline health worker community maternal and newborn health training. Midwifery 2013;29(3):251-259 [FREE Full text] [doi: 10.1016/j.midw.2012.01.006] [Medline: 22342173]
13. Krishna S, Boren SA, Balas EA. Healthcare via cell phones: a systematic review. Telemed J E Health 2009;15(3):231-240 [FREE Full text] [doi: 10.1089/tmj.2008.0099] [Medline: 19382860]
14. Free C, Phillips G, Galli L, Watson L, Felix L, Edwards P, et al. The effectiveness of mobile-health technology-based health behaviour change or disease management interventions for health care consumers: a systematic review. PLoS Med 2013;10(1):e1001362 [FREE Full text] [doi: 10.1371/journal.pmed.1001362] [Medline: 23349621]
15. Lim MSC, Hocking JS, Hellard ME, Aitken CK. SMS STI: a review of the uses of mobile phone text messaging in sexual health. Int J STD AIDS 2008;19(5):287-290 [FREE Full text] [doi: 10.1258/ijsa.2007.007264] [Medline: 18482956]
16. Fjeldsoe BS, Marshall AL, Miller YD. Behavior change interventions delivered by mobile telephone short-message service. Am J Prev Med 2009;36(2):165-173 [FREE Full text] [doi: 10.1016/j.amepre.2008.09.040] [Medline: 19135907]
17. Pop-Eleches C, Thirumurthy H, Habyarimana JP, Zivin JG, Goldstein MP, de Walque D, et al. Mobile phone technologies improve adherence to antiretroviral treatment in a resource-limited setting: a randomized controlled trial of text message reminders. AIDS 2011;25(6):825-834 [FREE Full text] [doi: 10.1097/QAD.0b013e32834380c1] [Medline: 21252632]
18. Horvath T, Azman H, Kennedy GE, Rutherford GW. Mobile phone text messaging for promoting adherence to antiretroviral therapy in patients with HIV infection. Cochrane Database Syst Rev 2012;2012(3):CD009756 [FREE Full text] [doi: 10.1002/14651858.CD009756] [Medline: 22419345]
19. Lester RT, Ritvo P, Mills EJ, Kariri A, Karanja S, Chung MH, et al. Effects of a mobile phone short message service on antiretroviral treatment adherence in Kenya (WelTel Kenya1): a randomised trial. Lancet 2010;376(9755):1838-1845 [FREE Full text] [doi: 10.1016/S0140-6736(10)61997-6] [Medline: 21071074]
20. Vital Wave Consulting. mHealth for development: the opportunity of mobile technology for healthcare in the developing world. United Nations Foundation. Washington DcBerkshire, UK; 2009. URL: https://reliefweb.int/report/world/mhealth-development-opportunity-mobile-technology-healthcare-developing-world [accessed 2024-02-15]
21. MHealth: New Horizons for Health Through Mobile Technologies. Geneva: World Health Organization; 2011.
22. Marshall C, Lewis D, Whittaker M. mHealth Technologies in Developing Countries: a Feasibility Assessment and a Proposed Framework. Herston, Australia: University of Queensland; 2013:1-47.
23. Shiferaw KB, Mehari EA. Modeling predictors of acceptance and use of electronic medical record system in a resource limited setting: using modified UTAUT model. Inform Med Unlocked 2019;17:100182 [FREE Full text] [doi: 10.1016/j.imu.2019.100182]
24. Andreatta P, Debpuur D, Danquah A, Perosky J. Using cell phones to collect postpartum hemorrhage outcome data in rural Ghana. Int J Gynaecol Obstet 2011;113(2):148-151 [FREE Full text] [doi: 10.1016/j.ijgo.2010.11.020] [Medline: 21420086]
25. Chang LW, Kagaayi J, Arem H, Nakigozi G, Ssempijja V, Serwadda D, et al. Impact of a mHealth intervention for peer health workers on AIDS care in rural Uganda: a mixed methods evaluation of a cluster-randomized trial. AIDS Behav 2011;15(8):1776-1784 [FREE Full text] [doi: 10.1007/s10461-011-9995-x] [Medline: 21739286]
26. Wills MJ, El-Gayar OF, Bennett D. Examining healthcare professionals' acceptance of electronic medical records using UTAUT. Issues Inf Syst 2008;9(2):396-401 [FREE Full text]
27. Arman AA, Hartati S. Development of user acceptance model for electronic medical record system. : IEEE; 2015 Presented at: 2015 International Conference on Information Technology Systems and Innovation (ICITSI); November 16-19, 2015; Bandung, Indonesia. [doi: 10.1109/icitsi.2015.7437724]
28. Cocosila M, Archer N. A blended model of electronic medical record system adoption in Canadian medical practices. Commun Assoc Inf Syst 2016;39(1):483-508 [FREE Full text] [doi: 10.17705/1cais.03922]
29. Hung SY, Yu WJ, Tsai MF, Yen DC. What drives physicians' intention to use electronic medical record system: the roles of perceived service level, computer self-efficacy, and perceived risk. 2013 Presented at: Pacific Asia Conference on Information Systems 2013 Proceedings; June 18, 2013; Virtual p. 194.
30. Venugopal P, Jinka S, Priya SA. User acceptance of electronic health records: cross validation of UTAUT model. Glob Manag Rev 2016;10(3):42-54 [FREE Full text]
31. Tubaishat A. Perceived usefulness and perceived ease of use of electronic health records among nurses: application of technology acceptance model. Inform Health Soc Care 2018;43(4):379-389 [FREE Full text] [doi: 10.1080/17538157.2017.1363761] [Medline: 28920708]
32. Medhanyie AA, Little A, Yebyo H, Spigt M, Tadesse K, Blanco R, et al. Health workers' experiences, barriers, preferences and motivating factors in using mHealth forms in Ethiopia. Hum Resour Health 2015;13(1):2 [FREE Full text] [doi: 10.1186/1478-4491-13-2] [Medline: 25588973]

XSL•FO
RenderX

33. Little A, Medhanyie A, Yebyo H, Spigt M, Dinant GJ, Blanco R. Meeting community health worker needs for maternal health care service delivery using appropriate mobile technologies in Ethiopia. PLoS One 2013;8(10):e77563 [FREE Full text] [doi: 10.1371/journal.pone.0077563] [Medline: 24204872]
34. Florez-Arango JF, Iyengar MS, Dunn K, Zhang J. Performance factors of mobile rich media job aids for community health workers. J Am Med Inform Assoc 2011;18(2):131-137 [FREE Full text] [doi: 10.1136/jamia.2010.010025] [Medline: 21292702]
35. Mushamiri I, Luo C, Iiams-Hauser C, Amor YB. Evaluation of the impact of a mobile health system on adherence to antenatal and postnatal care and prevention of mother-to-child transmission of HIV programs in Kenya. BMC Public Health 2015;15(1):102 [FREE Full text] [doi: 10.1186/s12889-015-1358-5] [Medline: 25886279]
36. The mobile economy: Sub-Saharan Africa 2014. GSMA Intelligence. 2014. URL: https://www.gsma.com/publicpolicy/wp-content/uploads/2015/02/The-Mobile-Economy-Sub-Saharan-Africa-2014-1-pager-FINAL.pdf [accessed 2024-02-15]
37. Mendoza G, Okoko L, Morgan G, Konopka S. mHealth compendium. Volume Two. United States Agency for International Development. 2013. URL: https://msh.org/wp-content/uploads/2013/09/usaid_mhealth_compendium_vol._2_us_letter_final_desktopprint_with_infogr.pdf [accessed 2024-02-15]
38. Klasnja P, Pratt W. Healthcare in the pocket: mapping the space of mobile-phone health interventions. J Biomed Inform 2012;45(1):184-198 [FREE Full text] [doi: 10.1016/j.jbi.2011.08.017] [Medline: 21925288]
39. Ahlan AR, Ahmad BI. User acceptance of health information technology (HIT) in developing countries: a conceptual model. Procedia Technol 2014;16:1287-1296 [FREE Full text] [doi: 10.1016/j.protcy.2014.10.145]
40. Al-Aswad AM, Brownsell S, Palmer R, Nichol JP. A review paper of the current status of electronic health records adoption worldwide: the gap between developed and developing countries. J Health Inform Dev Ctries 2013;7(2):153-164 [FREE Full text]
41. Boonstra A, Broekhuis M. Barriers to the acceptance of electronic medical records by physicians from systematic review to taxonomy and interventions. BMC Health Serv Res 2010;10(1):231 [FREE Full text] [doi: 10.1186/1472-6963-10-231] [Medline: 20691097]
42. Shiells K, Holmerova I, Steffl M, Stepankova O. Electronic patient records as a tool to facilitate care provision in nursing homes: an integrative review. Inform Health Soc Care 2019;44(3):262-277 [FREE Full text] [doi: 10.1080/17538157.2018.1496091] [Medline: 30125140]
43. Gao S, Krogstie J, Siau K. Developing an instrument to measure the adoption of mobile services. Mob Inf Syst 2011;7(1):45-67 [FREE Full text] [doi: 10.1155/2011/831018]
44. Venkatesh V, Davis FD, Morris MG. Dead or alive? the development, trajectory and future of technology adoption research. J Assoc Inf Syst 2007;8(4):267-286 [FREE Full text] [doi: 10.17705/1jais.00120]
45. Venkatesh V, Morris MG, Davis GB, Davis FD. User acceptance of information technology: toward a unified view. MIS Quarterly 2003;27(3):425-478 [FREE Full text] [doi: 10.2307/30036540]
46. Venkatesh V, Thong JYL, Xu X. Consumer acceptance and use of information technology: extending the unified theory of acceptance and use of technology. MIS Quarterly 2012;36(1):157-178 [FREE Full text] [doi: 10.2307/41410412]
47. Davis FD. Perceived usefulness, perceived ease of use, and user acceptance of information technology. MIS Quarterly 1989;13(3):319-340 [FREE Full text] [doi: 10.2307/249008]
48. Ryan RM, Deci EL. Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. Am Psychol 2000;55(1):68-78. [doi: 10.1037//0003-066x.55.1.68] [Medline: 11392867]
49. Limayem M, Hirt SG, Cheung CMK. How habit limits the predictive power of intention: the case of information systems continuance. MIS Quarterly 2007;31(4):705-737 [FREE Full text] [doi: 10.2307/25148817]
50. Aarts H, Verplanken B, van Knippenberg A. Predicting behavior from actions in the past: repeated decision making or a matter of habit? J Applied Social Pyschol 2006;28(15):1355-1374. [doi: 10.1111/j.1559-1816.1998.tb01681.x]
51. Matsunaga M. Item parceling in structural equation modeling: a primer. Commun Methods Meas 2008;2(4):260-293 [FREE Full text] [doi: 10.1080/19312450802458935]
52. Soliz J, Harwood J. Shared family identity, age salience, and intergroup contact: investigation of the grandparent–grandchild relationship. Commun Monogr 2006;73(1):87-107 [FREE Full text] [doi: 10.1080/03637750500534388]
53. Rogers WM, Schmitt N. Parameter recovery and model fit using multidimensional composites: a comparison of four empirical parceling algorithms. Multivariate Behav Res 2004;39(3):379-412 [FREE Full text] [doi: 10.1207/s15327906mbr3903_1]
54. Bentler PM. Comparative fit indexes in structural models. Psychol Bull 1990;107(2):238-246. [doi: 10.1037/0033-2909.107.2.238] [Medline: 2320703]
55. Chau PYK. Reexamining a model for evaluating information center success using a structural equation modeling approach. Decis Sci 1997;28:309-334 [FREE Full text]
56. Browne MW, Cudeck R. Alternative ways of assessing model fit. In: Bollen KA, Long JS, editors. Testing Structural Equation Models. Newbury Park, CA: Sage Publications; 1993:136-162.
57. Bentler PM, Bonett DG. Significance tests and goodness of fit in the analysis of covariance structures. Psychol Bull 1980;88(3):588-606. [doi: 10.1037/0033-2909.88.3.588]

XSL•FO
RenderX

58. Odendaal WA, Watkins JA, Leon N, Goudge J, Griffiths F, Tomlinson M, et al. Health workers' perceptions and experiences of using mHealth technologies to deliver primary healthcare services: a qualitative evidence synthesis. Cochrane Database Syst Rev 2020;3(3):CD011942 [FREE Full text] [doi: 10.1002/14651858.CD011942.pub2] [Medline: 32216074]
59. Cansdale LG, Kelly G, Khashan A, Malata A, Kachale F, Lissauer D, et al. Use of mHealth tools to register birth outcomes in low-income and middle-income countries: a scoping review. BMJ Open 2022;12(10):e063886 [FREE Full text] [doi: 10.1136/bmjopen-2022-063886] [Medline: 36223965]
60. Ballard M, Montgomery P. Systematic review of interventions for improving the performance of community health workers in low-income and middle-income countries. BMJ Open 2017;7(10):e014216 [FREE Full text] [doi: 10.1136/bmjopen-2016-014216] [Medline: 29074507]
61. Ahmed MH, Bogale AD, Tilahun B, Kalayou MH, Klein J, Mengiste SA, et al. Intention to use electronic medical record and its predictors among health care providers at referral hospitals, North-West Ethiopia, 2019: using unified theory of acceptance and use technology 2(UTAUT2) model. BMC Med Inform Decis Mak 2020;20(1):207 [FREE Full text] [doi: 10.1186/s12911-020-01222-x] [Medline: 32883267]
62. Kim S, Lee KH, Hwang H, Yoo S. Analysis of the factors influencing healthcare professionals' adoption of mobile Electronic Medical Record (EMR) using the Unified Theory of Acceptance and Use of Technology (UTAUT) in a tertiary hospital. BMC Med Inform Decis Mak 2016;16(1):12 [FREE Full text] [doi: 10.1186/s12911-016-0249-8] [Medline: 26831123]
63. Ouattara A. Antecedents of Employees' Behavioral Intentions Regarding Information Technology Consumerization. Minneapolis: Walden University; 2017.
64. Garavand A, Mohseni M, Asadi H, Etemadi M, Moradi-Joo M, Moosavi A. Factors influencing the adoption of health information technologies: a systematic review. Electron Physician 2016;8(8):2713-2718 [FREE Full text] [doi: 10.19082/2713] [Medline: 27757179]
65. Kuek A, Hakkennes S. Healthcare staff digital literacy levels and their attitudes towards information systems. Health Informatics J 2020;26(1):592-612 [FREE Full text] [doi: 10.1177/1460458219839613] [Medline: 30983476]
66. Ramachandran D, Canny J, Das PD, Cutrell E. Mobile-izing health workers in rural India. 2010 Presented at: CHI '10: CHI Conference on Human Factors in Computing Systems; April 10-15, 2010; Atlanta Georgia USA. [doi: 10.1145/1753326.1753610]
67. DeRenzi B, Findlater L, Payne J, Birnbaum B, Mangilima J, Parikh T, et al. Improving community health worker performance through automated SMS. 2012 Presented at: ICTD '12: Fifth International Conference on Information and Communication Technologies and Development; March 12-15, 2012; Atlanta Georgia USA. [doi: 10.1145/2160673.2160677]
68. Bennani AE, Oumlil R. IT acceptance by nurses in Morocco: application of a modified unified theory of acceptance and use of technology. IBIMA Bus Rev 2014;2014:849383 [FREE Full text] [doi: 10.5171/2014.849383]
69. Tavares J, Oliveira T. Electronic health record portal adoption: a cross country analysis. BMC Med Inform Decis Mak 2017;17(1):97 [FREE Full text] [doi: 10.1186/s12911-017-0482-9] [Medline: 28679423]
70. Mukred M, Yusof ZM, Alotaibi FM, Mokhtar UA, Fauzi F. The key factors in adopting an Electronic Records Management System (ERMS) in the educational sector: a UTAUT-based framework. IEEE Access 2019;7:35963-35980 [FREE Full text] [doi: 10.1109/access.2019.2904617]
71. Huang WM, Chen T, Hsieh CW. An empirical study on thephysicians' behavioral intentionwith electronic medical recordsystems in Taiwan. 2014 Presented at: Pacific Asia Conference on Information Systems 2014 Proceedings; June 01, 2014; Virtual p. 160.
72. Gagnon MP, Ghandour EK, Talla PK, Simonyan D, Godin G, Labrecque M, et al. Electronic health record acceptance by physicians: testing an integrated theoretical model. J Biomed Inform 2014;48:17-27 [FREE Full text] [doi: 10.1016/j.jbi.2013.10.010] [Medline: 24184678]
73. Agarwal S, Perry HB, Long LA, Labrique AB. Evidence on feasibility and effective use of mHealth strategies by frontline health workers in developing countries: systematic review. Trop Med Int Health 2015;20(8):1003-1014 [FREE Full text] [doi: 10.1111/tmi.12525] [Medline: 25881735]
74. Hampshire K, Mwase-Vuma T, Alemu K, Abane A, Munthali A, Awoke T, et al. Informal mhealth at scale in Africa: opportunities and challenges. World Dev 2021;140:105257 [FREE Full text] [doi: 10.1016/j.worlddev.2020.105257] [Medline: 33814676]
75. Cocosila M, Archer N. Adoption of mobile ICT for health promotion: an empirical investigation. Electron Markets 2010;20(3):241-250 [FREE Full text] [doi: 10.1007/s12525-010-0042-y]
76. WHO Guideline. Recommendations on digital interventions for health system strengthening. World Health Organization. 2019. URL: https://www.who.int/publications/i/item/9789241550505 [accessed 2024-02-15]
77. Braun R, Catalani C, Wimbush J, Israelski D. Community health workers and mobile technology: a systematic review of the literature. PLoS One 2013;8(6):e65772 [FREE Full text] [doi: 10.1371/journal.pone.0065772] [Medline: 23776544]
78. Gustafson DH, Sainfort F, Eichler M, Adams L, Bisognano M, Steudel H. Developing and testing a model to predict outcomes of organizational change. Health Serv Res 2003;38(2):751-776 [FREE Full text] [doi: 10.1111/1475-6773.00143] [Medline: 12785571]

XSL•FO
RenderX

79. Thondoo M, Strachan DL, Nakirunda M, Ndima S, Muiambo A, Källander K, et al. Potential roles of mhealth for community health workers: formative research with end users in Uganda and Mozambique. JMIR Mhealth Uhealth 2015;3(3):e76 [FREE Full text] [doi: 10.2196/mhealth.4208] [Medline: 26206419]
80. White A, Thomas DSK, Ezeanochie N, Bull S. Health worker mHealth utilization: a systematic review. Comput Inform Nurs 2016;34(5):206-213 [FREE Full text] [doi: 10.1097/CIN.0000000000000231] [Medline: 26955009]
81. Russpatrick S, Sæbø J, Romedenne M, Dubent L. The state of community health information systems in West and Central Africa. J Glob Health Rep 2019;3:e2019047 [FREE Full text] [doi: 10.29392//001c.12020]
82. Jeremie N, Kaseje D, Olayo R, Akinyi C. Utilization of community-based health information systems in decision making and health action in Nyalenda, Kisumu County, Kenya. Univers J Med Sci 2014;2(4):37-42 [FREE Full text] [doi: 10.13189/ujmsj.2014.020401]

Abbreviations

CHW: community health worker
eCHIS: electronic community health information system
HEP: Health Extension Program
HEW: health extension worker
mHealth: mobile health
UTAUT: Unified Theory of Acceptance and Use of Technology
UTAUT2: Unified Theory of Acceptance and Use of Technology 2

Edited by A Kushniruk; submitted 07.03.23; peer-reviewed by C Urquhart, D Carvalho, DEE A Jalil; comments to author 22.07.23; revised version received 31.08.23; accepted 05.10.23; published 04.03.24.

Please cite as:
Hailemariam T, Atnafu A, Gezie L, Kaasbøll J, Klein J, Tilahun B
Intention to Use an Electronic Community Health Information System Among Health Extension Workers in Rural Northwest Ethiopia: Cross-Sectional Study Using the Unified Theory of Acceptance and Use of Technology 2 Model
JMIR Hum Factors 2024;11:e47081
URL: https://humanfactors.jmir.org/2024/1/e47081
doi:10.2196/47081
PMID:38437008

©Tesfahun Hailemariam, Asmamaw Atnafu, Lemma Gezie, Jens Kaasbøll, Jorn Klein, Binyam Tilahun. Originally published in JMIR Human Factors (https://humanfactors.jmir.org), 04.03.2024. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on https://humanfactors.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Exploring the Use of Persuasive System Design Principles to Enhance Medication Incident Reporting and Learning Systems: Scoping Reviews and Persuasive Design Assessment

Kiemute Oyibo[1], PhD; Paola A Gonzalez[2], PhD; Sarah Ejaz[1], BSc; Tasneem Naheyan[1], MSc; Carla Beaton[3], RPh; Denis O'Donnell[4], PharmD; James R Barker[2†], PhD

[1]Department of Electrical Engineering and Computer Science, Lassonde Research Centre, York University, North York, ON, Canada
[2]Faculty of Management, Dalhousie University, Halifax, NS, Canada
[3]Pharmapod, Think Research Corporation, Toronto, ON, Canada
[4]CareRx, Clinical Excellence, Toronto, ON, Canada
[†]deceased

Corresponding Author:
Kiemute Oyibo, PhD
Department of Electrical Engineering and Computer Science, Lassonde Research Centre
York University
4751 Keele Street
North York, ON, M3J 2N9
Canada
Phone: 1 416 736 5053
Email: kiemute.oyibo@yorku.ca

Abstract

Background: Medication incidents (MIs) causing harm to patients have far-reaching consequences for patients, pharmacists, public health, business practice, and governance policy. Medication Incident Reporting and Learning Systems (MIRLS) have been implemented to mitigate such incidents and promote continuous quality improvement in community pharmacies in Canada. They aim to collect and analyze MIs for the implementation of incident preventive strategies to increase safety in community pharmacy practice. However, this goal remains inhibited owing to the persistent barriers that pharmacies face when using these systems.

Objective: This study aims to investigate the harms caused by medication incidents and technological barriers to reporting and identify opportunities to incorporate persuasive design strategies in MIRLS to motivate reporting.

Methods: We conducted 2 scoping reviews to provide insights on the relationship between medication errors and patient harm and the information system–based barriers militating against reporting. Seven databases were searched in each scoping review, including PubMed, Public Health Database, ProQuest, Scopus, ACM Library, Global Health, and Google Scholar. Next, we analyzed one of the most widely used MIRLS in Canada using the Persuasive System Design (PSD) taxonomy—a framework for analyzing, designing, and evaluating persuasive systems. This framework applies behavioral theories from social psychology in the design of technology-based systems to motivate behavior change. Independent assessors familiar with MIRLS reported the degree of persuasion built into the system using the 4 categories of PSD strategies: primary task, dialogue, social, and credibility support.

Results: Overall, 17 articles were included in the first scoping review, and 1 article was included in the second scoping review. In the first review, significant or serious harm was the most frequent harm (11/17, 65%), followed by death or fatal harm (7/17, 41%). In the second review, the authors found that iterative design could improve the usability of an MIRLS; however, data security and validation of reports remained an issue to be addressed. Regarding the MIRLS that we assessed, participants considered most of the primary task, dialogue, and credibility support strategies in the PSD taxonomy as important and useful; however, they were not comfortable with some of the social strategies such as cooperation. We found that the assessed system supported a number of persuasive strategies from the PSD taxonomy; however, we identified additional strategies such as tunneling, simulation, suggestion, praise, reward, reminder, authority, and verifiability that could further enhance the perceived persuasiveness and value of the system.

XSL•FO
RenderX

Conclusions: MIRLS, equipped with persuasive features, can become powerful motivational tools to promote safer medication practices in community pharmacies. They have the potential to highlight the value of MI reporting and increase the readiness of pharmacists to report incidents. The proposed persuasive design guidelines can help system developers and community pharmacy managers realize more effective MIRLS.

(JMIR Hum Factors 2024;11:e41557) doi:10.2196/41557

KEYWORDS

medication incident; reporting system; persuasive technology; persuasive design; medication; persuasive system design; pharmacy; pharmaceutic; pharmacology; drug reporting; drug event; adverse event; incident management

Introduction

Overview

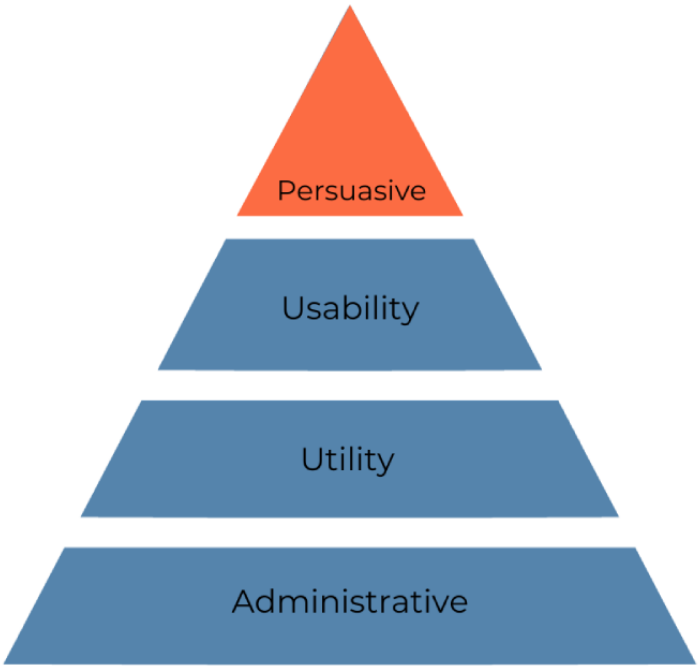
Medication errors are one of the leading causes of death in many countries worldwide [1,2]. For example, in the United States alone, 7000 to 9000 patients die annually owing to these errors. In Canada, where medical errors (labeled as the third leading cause of death after cancer and heart disease) account for 28,000 deaths annually, every minute and 18 seconds a patient gets harmed because of unintended errors, with medication errors being the most frequent [3]. Wrong medication (eg, because of similar naming, similar packaging, illegible handwriting, and incorrect drug selection) and wrong dose are among the most common medication errors in community pharmacies [3-5]. In particular, advanced drug preparation and administration without double checking [6] and heavy workflow [7] have been identified as key contributing factors to medication errors. However, there may be many more contributing underlying factors that go unreported by pharmacists and other health professionals. For example, a survey on medication administration errors among nurses in South Korea showed that 63.6% of the respondents had been involved in medication errors once or more in the previous month. However, only 28.3% of the participants reported the incidents [6]. Underreporting of medication errors, which is a global issue [8-11], has several implications bordering on shared learning, patient safety, and financial cost. In the United States, for example, psychological or physical pain and distress aside, “the total cost of looking after patients with medication-associated errors exceeds US $40 billion each year, with over 7 million patients affected” [4]. Moreover, underreporting of medication errors and incidents might limit individual and organizational learning from their occurrence [12,13].

The continuous evolution of pharmacotherapy and changing demands on the community pharmacy necessitate constant vigilance to detect new types of medication errors [14]. In a study among hospital pharmacists in South Korea, Hee-Jin et al [15] found that “five or more near misses per month were experienced by 14.8%, 4.3%, and 43.9% of respondents for dispensing, administration, and prescribing errors, respectively.” Moreover, research has shown that medication errors that lead to patient harm are common in medical care including community pharmacy [2,16-19]. Frequent reporting of all medication incidents (MIs) and near-miss events has the potential to improve patient safety through shared learning, which will enable the reduction of recurrence and prevention of MIs in the future [20,21]. Without adequate user reporting, none of the laudable objectives of reporting systems, including identification of gaps and resource development to support patient safety, can be achieved [7]. Medication error reporting is a common metric used to assess the quality of care provided by the health care system [21]. However, research has shown that employees are less motivated to report medication errors [22-25]. Hence, there is a need to find ways to motivate pharmacists and pharmacy technicians to report MIs more often to foster shared learning, prevention of recurrence, and patient safety. The question then is, *How can we motivate pharmacists and pharmacy technicians to report MIs more frequently using persuasive design principles embedded in digital technologies?*

Although some guiding principles have been proposed to alleviate the barriers to MI reporting, these principles, from a user experience (UX) design perspective, are not aimed at motivating pharmacists to report MIs regularly. From our literature search, we identified 4 categories of principles that can guide the design of Medication Incident Reporting and Learning Systems (MIRLS) to improve their adoption and usability. They include administrative principles, usability, utility principles, and persuasive design principles (Figure 1). Administrative principles refer to the organizational processes and policies implemented to enable and encourage employees to report medication errors regularly without fear of consequences. These principles form the basis of MIRLS, upon which the other categories of principles build. Usability and utility principles refer to the UX features that enable a user to report medication errors with ease, effectiveness, efficiency, and satisfaction [26]. Persuasive principles refer to the motivational affordances of a system that facilitate, nudge, and motivate a user to report medication errors. Current MIRLS mainly focus on the administrative, usability and utility-based principles. Typical examples of administrative principles include voluntary use, anonymity, confidentiality, and nonpunitive consequences. Examples of usability and utility-based principles, particularly in the Think Research and Pharmapod system, include ease of use, use of a standard taxonomy, searchability, retrievability, report generation, and root cause analysis [7,14,27].

XSL•FO
RenderX

Figure 1. Four key design principles for Medication Incident Reporting and Learning Systems in community pharmacy.



Apart from the administrative, usability, and utility principles, we argue that persuasive design principles hold potential to increase MI reporting among pharmacists. Persuasive design principles embedded in digital technologies, also known as persuasive technology, can motivate increased reporting of MIs from community pharmacies, as research in other health domains has shown [28]. Hence, this study proposes the use of persuasive design principles, which build on the 3 other categories of principles (Figure 1), to motivate users of MIRLS to report incidents and near misses more often.

Using Think Research, also known as Pharmapod, a cloud-based MIRLS for reporting and reducing incidents in community pharmacies [29], as a case study, this study (1) assesses 1 MIRLS based on the Persuasive System Design (PSD) taxonomy proposed by Oinas-Kukkonen and Harjumaa [30] and (2) proposes persuasive design guidelines to help community pharmacy stakeholders at multiple levels (eg, facility, provincial, and national) integrate persuasive features into their MIRLS. The PSD taxonomy is a widely used framework in the persuasive technology domain for analyzing, designing, implementing, and evaluating persuasive systems. Persuasive strategies from the PSD taxonomy can enhance MIRLS, making them more effective in promoting patient safety and shared learnings among practitioners [31,32]. Moreover, the study presents a summary of the results on the relationship between medication errors and harm and the information system–based barriers to MI reporting to ground the research.

Background and Related Work

In this section, we present an overview of relevant studies on the relationship between medication errors and patient harm and the organizational and information system barriers to reporting.

Medication Errors and Patient Harm

Several studies have been conducted to investigate the prevalence, nature, severity, and effects of MIs. West et al [16] investigated the relationship between medical errors and patient harm in primary care. They found that clinical harm to patients was reported in >10% of the 608 primary care medical error reports, with prescription-related errors most frequently linked to clinical harm. Similarly, Robb et al [17] investigated the relationship between medication and patient harm in hospitals in New Zealand. The authors confirmed the findings of earlier studies that showed that medication-related harms were common in both hospitals and the community, posing a substantial burden for patients and the health care system. In particular, they found that 923 harms were identified among 751 patients, with 28% of them experiencing ≥1 of the medication-related harms. They also found that older and female patients and those who had an increased length of stay were more likely to be harmed. Moreover, 65% of the harms occurred during an inpatient stay and 29% originated from the community and resulted in an admission. Riordan et al [18] investigated discharge prescription errors and their propagation after the discharge of patients. They found that 43% of the patients included in the study experienced postdischarge medication errors, with 86% of them being at risk of moderate harm. Moreover, 88% of the errors were discharge prescription errors that persisted after the discharge.

Most recently, Alqenae et al [2] conducted a systematic review, which they regarded as the first, to explore the prevalence and nature of medication errors and adverse drug events after hospital discharge. The review found that the median rate of medication error was approximately 50% among adult and older patients after hospital discharge, with approximately 20% of the patients in the studies reported to be affected by adverse drug events (such as antibiotics, antidiabetics, analgesics, and cardiovascular drugs) after hospital discharge. Panagioti et al [19] conducted a systematic review and meta-analysis of the

XSL•FO
RenderX

prevalence, severity, and nature of preventable patient harm across a range of medical care settings. They found that 5% of the patients were exposed to preventable harm in medical care and 25% of the incidents, which are drug related, accounted for the largest proportion of preventable patient harm, with 12% of the preventable patient harms being severe or leading to death. They asserted that there are limited quality improvement practices specifically targeting incidents that cause preventable harm to patients. They added that designing and implementing evidence-based mitigation strategies specifically targeting preventable patient harm could lead to substantial service quality improvements that are cost effective. This conclusion by Panagioti et al [19], coupled with the prevalence of medication errors in community pharmacy, partly informs this conceptual paper aimed to incorporate persuasive principles in MIRLS to increase medication error reporting and patient safety.

Organizational Barriers to MI Reporting

Researchers have identified several organizational barriers (both administrative and personal) leading to underreporting of medication errors and incidents in community pharmacy [21]. In the long run, these barriers can adversely affect patient safety owing to lack of shared learning among pharmacists within and across organizations because of underreporting [12,13]. Key barriers include fear of consequences such as punitive and disciplinary actions, negative or lack of administrative feedback, poor work climate or culture, inadequate training, and time constraint (Textbox 1) [7,8,21]. For example, Bahadori et al [9] found that the most important reasons for not reporting medication errors were administrative factors including the process of reporting and fear of the consequences of reporting. Research has also shown that personal (ie, sociodemographic) factors can impact medication error reporting. For instance, Aljabari and Kadhim [8] found that younger and lesser experienced professionals and staff with shorter employment periods were less likely to report medication errors. We argue that administrative barriers (such as time constraint and high workload) and perceived low value of the reporting system could be mitigated by using persuasive technologies to facilitate and ensure convenient reporting of MIs and errors. For example, persuasive design features (such as reminders to complete saved draft reports, notifications about the utility and value of reporting, and encouraging messages) may facilitate MI reporting.

Textbox 1. Administrative barriers to reporting medication errors and incidents.

Fear of consequences

- Negative consequences such as blame, shame, professional reputation damage, relationship damage, loss of privileges, medical malpractice lawsuit, relief from certain duties, and loss of job [4,9,33,34].

Lack of feedback

- Lack of useful feedback or negative feedback from administrative teams, such as pharmacy managers, regarding previously reported medication errors [33,34].

Poor work climate or culture

- Blaming staff and not the system or culture, poor support system, poor teamwork, poor organizational leadership, and lack of confidentiality in handling reports [33,35].

Miscommunication

- Poor communication among staff or between staff and patients [36].

Inadequate training of staff

- Difficulty in using the reporting system, poor understanding of the importance or value of reporting, poor understanding of errors, lack of clear definition of incident or near miss, and lack of a well-defined protocol on what events need to be reported [21,35].

Time constraint

- Work pressure and the lack of budgeted time to properly report errors, especially in the midst of a busy work schedule and high workload resulting in lack of enough breaks [7,35,36].

Information System Barriers to Patient Safety

Research has identified technological barriers that hamper patient safety in different health information systems and domains [37-40]. The primary barrier among them is the usability and poor design of health information systems [41]. Ratwani et al [42] found across 3 health care institutions that the usability of electronic health records accounted for more than a third of medication errors in 9000 pediatric patient safety reports. Kushniruk et al [43] evaluated the usability of a handheld prescription writing application. They found various usability problems (most of which relate to interface design) and actual errors in entering prescription data. In particular, they found that certain types of usability problems such as display visibility and ergonomics-related wrong data entry were closely linked to the occurrence of specific types of errors in medication prescription. More recently, Adams et al [37] investigated the medication errors associated with health information technology use and the harm caused to the patient. They found that 55.85% (1508/2700) of the manually reviewed reports described a medication error associated with information technology use and 49.7% (750/1508) of these caused harm to

XSL•FO
RenderX

the patient. In particular, they found that 97.35% (1468/1508) of the medication errors associated with information technology were related to usability issues including data entry, workflow support, and alerting. On the basis of these findings, in the current MIRLS domain, we set out to uncover the information technology barriers that border on the usability and utility principles (Figure 1), which may lead to the low perceived value, utility, and use of MIRLS.

PSD of MIRLS

PSD was pioneered by Fogg [44] in the early 2000s in his seminal book, "Persuasive technology: using computers to change what we think and do." This entails the application of behavioral theories from social psychology in the design of technology-based systems to motivate behavior change. Hence, persuasive technology is defined as a motivational tool intentionally designed to change human attitudes and behaviors through persuasive techniques grounded in social psychology [44]. Fogg [44] first proposed a set of 7 persuasive strategies to motivate behavior change. Subsequently, Oinas-Kukkonen and Harjumaa [30] extended the list to 28 persuasive strategies, which are categorized into 4 functional groups (primary task support, dialogue support, social support, and system credibility support), each comprising 7 persuasive strategies. Oyibo [45] extended the primary task support and dialogue support groups with goal setting and verbal persuasion, respectively, increasing the total number of strategies in the PSD taxonomy to 30. The primary task support category, which includes tunneling, tailoring, and self-monitoring, is aimed at helping the user to perform a target behavior easily and effectively. The dialogue support category, which includes praise, reward, and suggestion, is aimed at motivating the user to perform the target behavior through feedback and dialogue with the persuasive system. The social support category, which includes social learning, social comparison, and competition, is aimed at motivating the user through social influence to perform the target behavior. Finally, the system credibility support category, which includes trustworthiness, surface credibility, and authority, is aimed at increasing the user's trust in the system by making the system look professional and credible [46].

Incorporating persuasive features into MIRLS has the potential to improve the rate of error reporting. St-Maurice et al [28] showed that, on average, the percentage of same-day data entries can be increased by 10% for each user by introducing new persuasive design features into a data entry system. On the basis of this prior research finding, we propose guidelines for incorporating persuasive design principles, drawn from the PSD taxonomy, into MIRLS using the Think Research or Pharmapod Incident Management (IM) system as a case study. The PSD taxonomy, which comprises 4 categories of persuasive strategies (primary task support, dialogue support, social support, and system credibility support), is a framework for analyzing, designing, implementing, and evaluating persuasive systems. A systematic review by Win et al [47] showed that primary task support and dialogue support are the most commonly used categories of persuasive strategies in medication management information systems. The review reported that tailoring, self-monitoring, and reminders, which belong to the primary task support category, are more likely to be implemented in medication management information systems than other persuasive strategies. In the case of MIRLS, the proposed persuasive strategy guidelines are aimed at enhancing system utility and facilitating the reporting of near misses and incidents. Research shows that the higher the perceived usefulness of health systems, the higher the number of users who find them more persuasive [48].

Methods

Overview

A total of 2 types of methods were used to address 3 research questions (RQs). They include scoping review and assessment of an existing MIRLS based on administrative, usability, utility, and persuasive features. The RQs are as follows:

1. RQ1. Is there a relationship between medication errors and patient harm?
2. RQ2. What are the information system–based barriers preventing pharmacists and pharmacy technicians from reporting medication errors?
3. RQ3. How can we motivate them to report MIs more frequently using persuasive design principles embedded in digital technologies?

Ethical Considerations

The assessment of our target system was aimed at quality improvement, thus ethical approval was not required [49,50].

Scoping Reviews

To address the first 2 RQs, the authors (KO, SE, and TN) conducted 2 scoping views in August 2023. The first review investigated the relationship between medication errors and patient harm in the pharmacy domain. The second review aimed to uncover usability and utility-related barriers to medication error reporting. We retrieved articles from 6 databases for each study, screened the articles, extracted the relevant data, and presented the results. For the first review, a total of 820 articles were retrieved from PubMed (n=41), Public Health Database (n=89), ProQuest (n=451), Scopus (n=97), ACM (n=42), and Global Health (n=22) using the search string: "(Medic* OR prescri* OR administ* OR drug*) AND (error* OR incident* OR accident* OR nearmiss* OR 'near miss*' OR mistake*) AND patient AND (harm* OR hurt* OR injur* OR wound* OR bruise* OR impairment* OR afflict*) AND pharmac*." A total of 215 duplicates were removed to arrive at 605 unique articles. These articles were screened based on title or abstract to arrive at 91 articles. Next, a full-text review was conducted to arrive at 14 included articles after excluding 77 ineligible articles. Finally, 3 more articles were included to the 14 through Google Scholar search, resulting in 17 articles for the final data analysis. For the second review, a total of 849 articles were retrieved from PubMed (n=268), Public Health Database (n=44), ProQuest (n=90), Scopus (n=448), ACM (n=10), and Global Health (n=45) using the search string: "(Medic* OR prescri* OR administ* OR drug*) AND (error* OR incident* OR accident* OR nearmiss* OR 'near miss*' OR mistake*) AND (report* OR submi* OR log*) AND (system* OR application* OR website* OR tool* OR platform* OR interface* OR technolog*) AND pharmac* AND (barrier* OR hinderance*

XSL•FO
RenderX

OR obstacle* OR drawback* OR setback* OR deterrent* OR limitation* OR shortcoming*." A total of 303 duplicates were removed to arrive at 546 unique articles. These articles were screened based on title or abstract to arrive at 12 articles. Upon the full-text review, we arrived at zero article for data extraction and analysis. Moreover, based on Google Scholar search, we found 1 article [13] that investigated the usability of MIRLS called the Medication Error Reporting App. However, this study did not investigate the relationship between the usability of the app and medication error.

Overview and Initial Assessment of an Existing MIRLS

The authors (KO and PAG) analyzed the Think Research or Pharmapod MIRLS, which is a cloud-based software platform for reporting medication errors (incidents and near misses). As stated on its website, Think Research or Pharmapod describes itself as "the first platform of its kind to pool and share patient safety data across borders, monitoring trends and causes behind medication errors, and empowering healthcare professionals locally to improve their practice" [29]. Our initial review of the system assessed it against the 3 key design principles shown in Figure 1. To assess the administrative and usability and utility principles, the first 2 authors went through the Think Research or Pharmapod system from one interface to another to elicit the supported principles. Next, we used the PSD taxonomy as an assessment framework and 3 assessors (study participants) to identify persuasive strategies fully or partially implemented in the Think Research or Pharmapod IM system. We first assessed the system to identify the existing persuasive strategies and then gathered data from 3 experienced users to propose opportunities for improvement. One of the authors, the vice president of the Quality Improvement and Innovations of Think Research or Pharmapod, arranged for 3 independent and experienced users of the Think Research or Pharmapod IM system from different pharmacies to assess the system against the PSD taxonomy and items. The first assessor was a pharmacist who had 1.5 years of experience using the system. The second assessor was a director in a health care company focused on patient and staff safety, with 1 year of experience working with the system. The third assessor was a senior technology manager at a leading Canadian pharmacy company, with 4 years of experience working with the system.

The authors (KO and PAG) asked the assessors to independently indicate whether each persuasive strategy in the PSD taxonomy is important or useful, present in the system or not, and where it could be found in the system. The implementation of each strategy from the PSD taxonomy was described to the participants in a tabular form. The participants independently responded to the questions and then came together to discuss and confirm their responses and resolve their differences with the first 2 authors. If at least 2 of the 3 assessors indicated or agreed that a given persuasive strategy is important and useful, "yes" is entered into the associated cell in the table, otherwise, "no." Similarly, if at least 2 assessors agreed that the strategy was present in the system (ie, said "yes"), "√" is entered into the cell associated with the status column. However, if ≥2 assessors agreed that the strategy was not present in the system (ie, said "no"), "X" is entered into the associated cell under the status column. Moreover, if at least 1 of the assessors agreed that the strategy was present in the system, but the implementation was limited, "√*" standing for "present but could be improved" is entered into the associated cell under the status column.

Results

In this section, we present the results of the scoping reviews and the initial assessment of the Think Research or Pharmapod IM system.

Scoping Reviews

Medication Errors and Patient Harm

In the first review, 41% (7/17) of the included articles originated from North America (United States [16,51-53], Canada [54,55], and Mexico [56]), 29% (5/17) from Europe (United Kingdom [12], Ireland [18], the Netherlands [57], Sweden [58], and Spain [59]), 23% (4/17) from Asia (Saudi Arabia [60,61], China [62], and Korea [63]), and 6% (1/17) from Oceania (New Zealand [17]). The articles were published between 2001 and 2023, with most of the articles (3/17, 18%) published in 2023. Most of the target populations were from North America (7/17, 41%), followed by Asia (5/17, 29%), Europe (4/17, 23%), and Oceania (1/17, 6%). Of the 17 articles, 1 (6%) each focused on target populations in Africa and South America. Table 1 shows 16 types of harms elicited from the included articles. These were caused by 59 types of medication errors such as wrong drugs, missing or wrong patient weight, prescription errors, dosing error, wrong or unclear dose or strength, wrong patient, and wrong duration, each of which was reported by at least 2 articles. Significant or serious harm was the most frequent harm; it was reported by 65% (11/17) of the articles, followed by death or fatal harm (7/17, 41%) and no harm or potential harm (4/17, 23%).

XSL•FO
RenderX

Table 1. Type or severity of harm caused by medication errors and the number of articles associated with them (N=17).

Type or severity of harm	Articles, n (%)
Significant or serious harm	11 (65)
Death or fatal harm	7 (41)
No harm or potential harm	4 (23)
Inconvenience	3 (18)
Adverse drug events	3 (18)
Mild harm	2 (12)
Moderate harm	2 (12)
Temporary injury or harm	2 (12)
Prolonged hospitalization	2 (12)
Life-threatening harm	2 (12)
Nonlife threatening	1 (6)
Risk to patient or others	1 (6)
Unstable situation	1 (6)
Unknown harm	1 (6)
Permanent harm	1 (6)
Intervention required	1 (6)

Information System Barriers to MI Reporting

One article [13] that investigated the usability of an MIRLS prototype called Medication Error Reporting App found that there was significant improvement in the mean usability score throughout the development process ($P<.001$). However, this mean improvement in usability did not impact the mean time to report medication errors using the app because the mean time was not significantly different between the phases of the development process. Overall, it was found that the testers including pharmacists found the app easy to use, but doctors and nurses were unfamiliar with the medication terms used, especially the medication process in which error occurred and type of error. More importantly, the authors reported that although testers might be willing to adopt the app to make reports in the future, they were apprehensive about data protection issues such as security and abuse of feedback featured in the app [13].

Initial Assessment of Existing MIRLS

In this section, we present the results that emanated from the initial assessment of the Think Research or Pharmapod IM system based on administrative, usability, utility, and persuasive principles.

Administrative and Usability and Utility Principle Support

The assessed system supported at least 75% (6/8) of the administrative guiding principles shown in Textbox 2, including voluntariness, anonymity, confidentiality of information, and nonpunitive measures. It also supported all 7 usability and utility-based principles, including ease of use, searchability and retrievability, standard taxonomy, report generation, and root cause analysis (Table 1).

XSL•FO
RenderX

Textbox 2. Items and questions asked of assessors.

Strategy code

- A codeword representing the persuasive strategy.

System capability

- A description of the persuasive strategy.

Important or useful

- An indication of the importance or usefulness of the strategy (yes or no).

Present in system

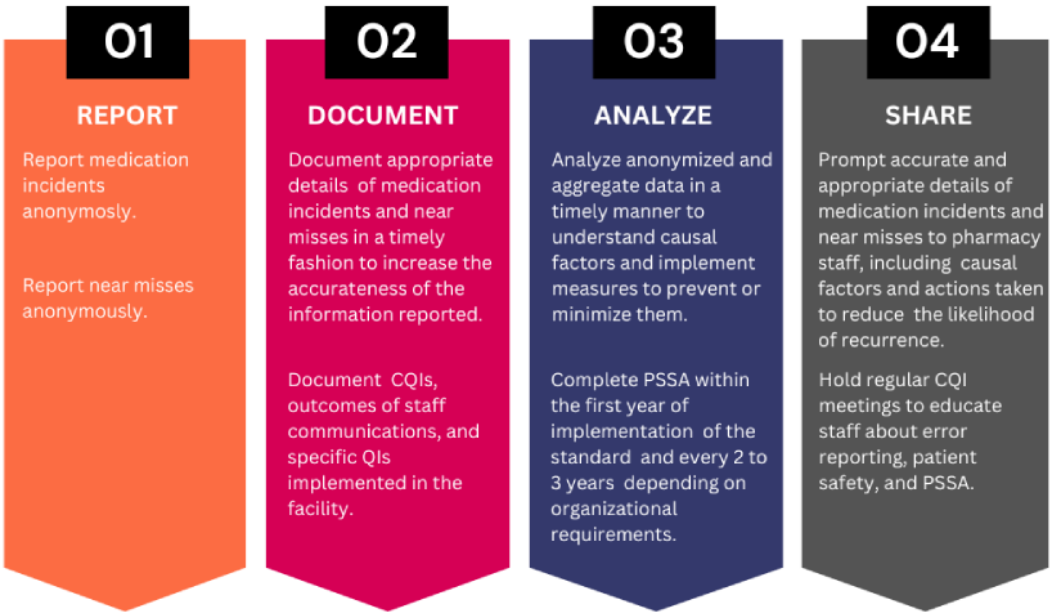
- An indication of the presence of the strategy in the system (yes or no).

Interface, tab, or comment

- Provision of the system interface or tab where the persuasive strategy can be found or a comment by the assessor.

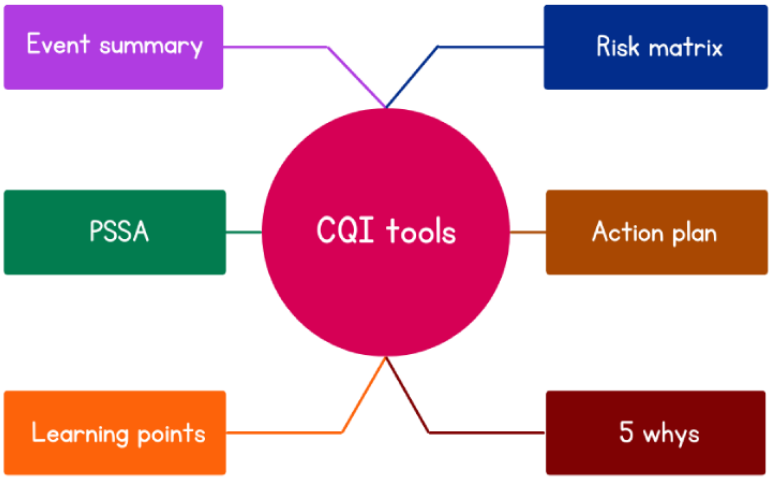
Moreover, the system promotes 4 key elements of patient medication safety: report, document, analyze, and share (Figure 2). The analyze and share elements are in addition supported by 6 main continuous quality improvement (CQI) tools. These tools are intended to foster patient safety in community pharmacy within a pharmacy team [20,64]. The tools include event summary, risk matrix, 5-whys template, action plan, learning points, and pharmacy safety self-assessment (Figure 3). An event summary is an incident and a root cause analysis tool. The risk matrix is a color-coded matrix that facilitates the assignment of a risk score based on the probability of recurrence of the incident or near miss at a specific severity level and its impact on a patient if it were to recur. The 5-whys is a tool that facilitates the analysis of an incident or near miss by answering the fundamental question, “Why did the incident occur?” 5 times. The 5-whys is a simple and well-recognized tool for determining the cause and effect of an incident objectively. Action plans is a tool to create and track smart actions of improvement. Learning points organizes identified gaps, for example, in workflows and processes and provides a means to share these learnings. Finally, the pharmacy safety self-assessment is a tool that allows the pharmacy team to proactively identify risks that may compromise patient safety and implement safe medication measures to address them [64].

Figure 2. Four key elements of patient medication safety.



XSL•FO
RenderX

Figure 3. Six continuous quality improvement (CQI) tools in the studied system aimed at fostering fundamental change among pharmacy team members. PSSA: pharmacy safety self-assessment.



Persuasive Principle Support

Tables 2-5 show the results of the assessment of the Think Research or Pharmapod IM system based on the primary, dialogue, social, and credibility support categories of the PSD taxonomy, respectively. The first column of Table 2 captures the coded name of the strategy and its description, the second column describes a yes or no response on the importance and usefulness of the strategy, and the third column describes a yes or no response on the presence of the strategy in the system (ie, status). A fourth column was also provided for the assessors to comment on the assessment of each strategy, for example, the location of the strategy in the system.

XSL•FO
RenderX

Table 2. Guidelines for incorporating the primary task support principles into Medication Incident Reporting and Learning Systems.

Strategy and implementation	I or U[a]	Status
Reduction		
Break down the medication incident and near-miss reporting process into a few simple steps to facilitate reporting [65].	Yes	✓[b]
Tunneling		
Guide the user through the reporting process in a step-by-step fashion, just as a software installation wizard [47].	Yes	X[c]
Goal setting		
Allow the user to set a goal, for example, minimum number of errors or CQI[d] reports to be submitted over a given period such as a week or month.	Yes	X
Self-monitoring		
Allow the user to track their progress after setting a report-based goal or when submitting a report, for example, through the display of a progress bar.	Yes	X
Allow the user to view their number of completed and uncompleted reports and averages per week, month, or year (eg, on their dashboard).	Yes	✓
Allow the user to track the levels of usefulness of their reports (eg, CQI, incident, or near miss) to others, for example, other users or colleagues "like" their anonymous reports as obtainable in YouTube and Facebook.	Yes	X
Tailoring		
Tailor what the user sees (eg, user profile, chart, content, and information) using group-based characteristics such as work experience and designation or role.	Yes	✓*[e]
Personalization		
Personalize the system (eg, information, report, and reminder) based on their interaction, for example, letting the user know where they left off or reminding them about incomplete tasks when they log in [7,66].	Yes	X
Customization		
Allow the user to customize the system (eg, profile, chart, content, information, and reminder) to suit their needs and preferences [66].	Yes	✓*
Simulation		
Show the user a cause-and-effect relationship of the benefit of incident or near miss or CQI reporting, for example, a study chart showing the higher the incidents reported, the lower the number of recurrences.	Yes	X
Rehearsal		
Provide a new user with a simulated environment to rehearse before making an actual report relating to an incident, near miss, or CQI.	Yes	✓*
Provide a new user with video tutorials on how to report a medication incident or near miss.	Yes	✓

[a]I or U: important or useful.

[b]Currently implemented.

[c]Not currently implemented.

[d]CQI: continuous quality improvement.

[e]Partially implemented and could be improved.

XSL•FO
RenderX

Table 3. Guidelines for incorporating dialogue support principles into Medication Incident Reporting and Learning Systems.

Strategy and implementation	I or U[a]	Status
Praise		
As a show of appreciation, praise or congratulate the user for submitting a near-miss or incident or CQI[b] report or for reaching a milestone using textual, visual, or audio-based feedback messages [67].	Yes	X[c]
Reward		
Reward the user with points, badges, etc, when they submit a report (early), achieve a goal or milestone, or others find their report useful (eg, by liking it), etc.	Yes	X
Allow the user over time to grow in the value of their contribution to the community. This can be based on the number, frequency, quality, earliness, and usefulness of their reports (to others), for example, from a silver to a gold valuable contributor of the community.	Yes	X
Reward the user for reporting or sharing action plans that improved safety in the pharmacy.	Yes	X
Reward the user for reporting positive experiences that led to improved safety in the pharmacy, for example, "good news" stories in addition to the negative "error" reports.	Yes	X
Suggestion		
Suggest to the user from time to time based on their profile, role, or interaction with the system new reports that may be interesting and beneficial to their practice [65].	Yes	X
Suggest to the user ways, processes, or methods through which others in the community prevent or address recurrence of certain near misses and incidents.	Yes	X
Suggest to the user standard, process-based solutions (eg, from the user's pharmacy, province, or professional organization) for addressing certain types of recurring incidents and near misses [65].	Yes	X
Provide the user with a list of "high-alert" medications or types of incidents that occur most often or require extra precautions and suggest best practices to reduce incidents and near misses associated with them [68].	Yes	✓*[d]
Feedback[e]		
Provide the user with summary feedback on their progress toward reaching their monthly, quarterly, or yearly goal (eg, "You have achieved 30% of your goal").	Yes	X
Provide the user with summary feedback on the usefulness of their reports to others (eg, "5% of the system users in the province [nation] found your report helpful").	Yes	X
Provide the user monthly, quarterly, or yearly summary feedback highlighting the most recurring types of near misses and incidents (eg, "Poor drug naming caused 5% of the near misses last year") [65].	Yes	✓[f]
Reminder		
Remind the user from time to time (eg, based on self-set goals) about the need to report near misses and incidents and about the benefits to other users and patient safety.	Yes	X
Remind the user from time to time to complete their CQI action plan that they have started.	Yes	X
Verbal persuasion[e]		
Allow management such as pharmacy managers and supervisors through personally sent messages to encourage users from time to time to report near misses and incidents, for example, "Alice, remember to report your near misses and incidents to improve patient safety. Yes, you can!"	N/A[g]	N/A
Emotional appeal[e]		
Use motivational messages to encourage users to report errors, for example, "To err is human, to share is divine" [69].	Yes	X
Liking		
Make the system to be visually attractive, for example, by using visually pleasing or appropriate colors to present charts, content, and important information.	Yes	✓*

[a]I or U: important or useful.

[b]CQI: continuous quality improvement.

[c]Not currently implemented.

[d]Partially implemented and could be improved.

[e]Not originally listed in the Persuasive System Design taxonomy.

[f]Currently implemented.

XSL•FO
RenderX

[g]N/A: not applicable.

Table 4. Guidelines for incorporating social support principles into Medication Incident Reporting and Learning Systems.

Strategy and implementation	I or U[a]	Status
Social learning		
Notify the user by email when other anonymous users submit an incident report (eg, containing the key points) or CQI[b] report that may be of interest to the user, just as in ResearchGate, for example, "John [a pseudonym], here's a new report we think you'll be interested in."	Yes	X[c]
Support chat room and discussion room to foster social support and shared learning [47]. This room can be anonymous.	Yes	X
Support a newsfeed (eg, as in Facebook) to highlight important reports the user may find useful and foster shared learning.	Yes	X
Social comparison		
Allow the user to compare their weekly, monthly, quarterly, or yearly reports with others, maintaining confidentiality (eg, at the city, zone, provincial, or national level).	Yes	✓*[d]
Competition		
Allow the user to see where they are compared with other anonymous (eg, on a leaderboard) at the pharmacy, provincial, or national level based on the total number, frequency, quality, or usefulness of their report to others (eg, over a weekly, monthly, or yearly period).	Yes	X
Cooperation		
Provide users the choice of being paired with another anonymous user, with the goal of motivating one another to achieve individual or collective goals.	No	X
Normative influence		
Inform users about the number of other anonymous users in the pharmacy, province, or nation that are reporting errors in a given period (eg, "10 other people submitted their incident reports today").	Yes	X
Social facilitation		
Make users, who are logged onto the system know that there are other anonymous users elsewhere (eg, in the facility, province, and nation), who are submitting or just submitted a report (eg, "5 other people are currently submitting their incident reports").	Yes	X
Social recognition		
Provide a means for committed user to be publicly recognized for being one of the "most valuable players" of the month, quarter, or year at the pharmacy, provincial, or national level based on certain criteria (eg, number, frequency, quality, or usefulness of their reports to the community).	Yes	X
Allow other users to rate users' reports anonymously based on how useful or helpful it is to them.	No	X

[a]I or U: important or useful.

[b]CQI: continuous quality improvement.

[c]Not currently implemented.

[d]Partially implemented and could be improved.

XSL•FO
RenderX

Table 5. Guidelines for incorporating system credibility support principles into Medication Incident Reporting and Learning Systems.

Strategy	Implementation	I or U[a]	Status
Authority	Present authority-based information and messages (eg, on the value of reporting incidents and near misses and the benefits it can have for the profession, staff, or patient safety) [47].	Yes	X[b]
Third-party endorsement	Demonstrate that the system is approved by authorities such as professional organizations, regulatory bodies, and government, for example, by displaying their corporate logos [65].	Yes	X
Expertise	The visual and functional design of the system should reflect professionalism, expertise, and be up to date to motivate users to use it.	Yes	✓*[c]
Trustworthiness	Build trust into the system, for example, by fostering anonymity, data aggregation, and keeping promises such as it not being used as a punitive tool to hold users accountable [65].	Yes	✓*
Surface credibility	Build surface credibility into the system through its visual design, for example, by reducing advertisements and ensuring users enter accurate information using taxonomy-based predefined options, checklists, and drop-downs [47].	Yes	✓[d]
Verifiability	Ensure presented information and messages (eg, on the value of error reporting to the profession, staff, or patient safety) are verifiable, for example, through a link to authority-based websites such as Institute for Safe Medication Practices and World Health Organization.	Yes	X
Real-world feel	The design of the system should mimic the paper-based error reporting forms (eg, [70]) as closely as possible to reduce the cognitive effort required by a new user to make the transition [71].	Yes	✓

[a]I or U: important or useful.

[b]Not currently implemented.

[c]Partially implemented and could be improved.

[d]Currently implemented.

Primary support strategies facilitate the key behaviors promoted by the system, such as reporting. Dialogue support strategies enable users to interact, engage with, and receive feedback from the system through text-, image-, audio-, and video-based dialogue. Social support strategies motivate users through social influence. Finally, credibility support strategies enable users to trust and rely on the system. In summary, based on the assessors' responses, most of the persuasive strategies (29/31, 94%) in the extended PSD taxonomy were considered important or useful, with approximately one-third (14/29, 48%) of them identified as present in the current Think Research or Pharmapod system. Approximately 23% (7/31) and 26% (8/31) of the strategies were considered fully or partially implemented (although they could be improved), respectively. More than 50% (16/31) of the strategies were considered not implemented, with most of them falling under the social support category.

Discussion

We have presented the results of 2 scoping reviews and the initial assessment of the Think Research or Pharmapod system. The following sections discuss the results with a focus on the persuasive design guidelines shown in Tables 2-5, which can inform the persuasive design of future MIRLS.

Summary of Scoping Review Findings

Table 1 shows the types of harm uncovered in the first scoping review. More than half (11/17, 65%) of the included articles reported that medicated errors caused serious harm to patients. In particular, 60% (3/5) of the articles reported serious harm, and 40% (2/5) of the articles reported fatal harm or death caused by medication errors such as wrong dose, drug, patient, and ambulatory pump (eg, [58]). Prescription error [16,18,59,63], wrong drugs [12,58,63], and dosing error [58,59,63] were the most frequent medication errors. For example, in the study by Fyhr and Akselsson [58], most severe medication errors occurred during prescribing and transcribing by physicians. The findings are an indication that medication errors have the potential to cause serious harm to patients, including death; hence, there is a need for interventions aimed to reduce them and increase patient safety (eg, by increasing reporting and shared learning within and across organizations). Moreover, in the second review on the usability of MIRLS, George et al [13] found that an iterative design has the potential to improve the usability of an MIRLS. However, their study suggested that there is a need to address issues surrounding data security and report validation to increase user acceptance and use.

Summary of Administrative and Usability and Utility Assessment

Our assessment shows that the Think Research or Pharmapod system implemented most of the administrative, usability and utility-based principles shown in Textboxes 1, 3 [7,14], and 4 [7,14]. Prior studies advocate most of these principles as essential actions and capabilities aimed at improving incident reporting and shared learning [7,9,14,33-36]. An anonymous reporting, for example, can mitigate the punitive perceptions of incident reporting [20]. However, the system only partially supported persuasive design principles. Persuasive design principles are intended to complement the administrative, usability and utility-based principles by improving the UX and motivating users to see value in reporting MIs and completing the CQI and learning tool reports. Persuasive design may in

XSL•FO
RenderX

turn mitigate some of the persistent barriers identified in Textbox 1.

Textbox 3. Administrative guiding principles for designing Medication Incident Reporting and Learning Systems [7,14].

Voluntariness

- Medication reporting will be voluntary.

Inclusiveness

- Professionals and consumers will be encouraged to participate.

Aggregation

- The reporting system will support anonymity and aggregation.

Confidentiality

- The system will provide confidentiality of reported information.

No consequence

- The system will clearly define and support a nonpunitive approach to reporting.

Type of report

- The system will encourage reporting of both potential and actual incidents and near misses.

Feedback

- The system will provide feedback on incident analysis and timely recommendations.

Workflow alignment

- The system should fit with the users' workflow.

Textbox 4. Usability and utility-based guiding principles for designing Medication Incident Reporting and Learning Systems [7,14].

Usability

- The system will be easy to use and time efficient.

Format multiplicity

- The system will support both electronic and paper formats.

Taxonomy

- The system will support standard taxonomy.

Outcome severity

- The system will support levels of severity of outcomes.

Searchability and retrievability

- The system will support searchable and retrievable data.

Report generation

- The system will support report generation.

Root cause analysis

- The system will support root cause analysis.

Summary of Persuasive Design Assessment

In this section, we discuss the results of the system assessment and persuasive design guidelines for designing future MIRLS, taking each category of the PSD taxonomy at a time.

Primary Support Assessment and Guidelines

In the primary support category (Table 2), all the persuasive strategies were considered important or useful, whereas over

XSL•FO
RenderX

55% (5/9) of them were deemed partially or fully implemented by the system.

Reduction

Reduction, which is considered important and present in the system by the assessors, entails breaking down the performance of a complex behavior into a few steps. In the context of MI reporting, this means making the reporting process simple and easy to carry out by users. Reduction is vital to ensuring and facilitating the report of MIs and near misses given the relatively high workload health professionals such as pharmacists handle on a daily basis [65]. In the Think Research or Pharmapod system, for example, to speed up the reporting process, predefined fields and system design widgets such as drop-downs are used to enter information about prescribed drugs, what happened, contributing factors, and harm caused. A critical aspect in realizing the effectiveness of the implementation of this and other PSD guidelines is the fit of the MI reporting task into users' workflow to facilitate regular reporting [7]. However, this examination is beyond the scope of this conceptual study.

Tunneling

Similar to reduction, tunneling (aka guided persuasion) aims at motivating users to report MIs and near misses. The tunneling strategy, which can be likened to the process of installing software on a computer using an installation wizard [72], is used to walk the user through predetermined steps in a structured manner. Two of the assessors agreed that tunneling is important or useful but not present in the studied system, with 1 of them remarking, "the report has four sections, then the CQI has colour coded features but they do not tunnel you in any direction." Once an incident report is completed and saved, the incident analysis interface (an event summary page containing a variety of management tools to prevent the recurrence of similar events in the future) opens automatically. However, the system does not tunnel the user in a specific direction. The third assessor, however, did not find tunneling useful in this context and commented, "No this MIRLS is not like an installation wizard. We like the flexibility provided today." Hence, owing to the mixed responses, a more comprehensive study among a larger target audience is required to understand the perceived usefulness of tunneling.

Goal Setting

Related to the commitment principle proposed by Cialdini [73], goal setting is known as one of the cornerstones of persuasive systems [74]. According to the commitment principle, people are more likely to follow through with a behavior if they make a commitment in written or verbal form to perform the behavior [75]. Studies have shown that people, regardless of culture, are more motivated by the commitment principle than by the other 5 principles of persuasion proposed by Cialdini [76,77]. Goal setting is more likely to be effective if set goals are specific, measurable, achievable, relevant, and timebound (SMART) [77]. The assessors agreed that goal setting is important or useful in MI reporting. One participant thought that the feature was present in the system already. Here, the assessor meant the CQI action planning. In general, both goal setting and action planning are related. However, action planning is concerned with how set goals can be achieved [78]. In the studied system, the CQI actions tool captures both actions (which can be regarded as CQI goals) and action plans (eg, addressing gaps in workflows and processes) [64]. Although we can submit the system-supported CQI goals, it did not support incident reporting goals. Regarding the former, one of the assessors stated that the action plan tab in the system allows a free-form type (such as textboxes that allow the user to type in anything without restrictions). However, it "could be improved by adding prompts for SMART [plans] to guide the user to complete [them] correctly. These action plans are incident specific. They do not allow overall SMART goals around frequency and quality of reporting. [Although], [t]here are dashboards of measurements, they do not include goals or thresholds as a comparison or guide."

Self-Monitoring

Self-monitoring goes hand in hand with goal setting in most implementations [78,79]. In other words, users should be able to visualize their progress toward the realization of their set goals. Self-monitoring is one of the cornerstones of persuasive systems [78] and one of the most requested persuasive features in health apps such as fitness apps [45]. In a systematic review, Matthews et al [80] found that 70% of the included articles evaluated physical activity apps that supported self-monitoring as a persuasive feature to motivate behavior change. Self-monitoring fosters self-reflection and raises users' consciousness of their responsibilities, which culminates in self-regulation and behavior change [78,81,82]. Self-monitoring can be compared with holding a mirror up to the user's face, and if the user does not like what they see, they do something about it. In work environments, employees' engagement in self-monitoring is considered a prerequisite for professional development [82]. In the studied system, self-monitoring is implemented in the form of incident and near-miss reports at the pharmacy, province, or national level. In the data warehouse interface, users can view the number of cases (incidents and near misses); number of events by harm levels, top 5 drugs; and what, why, and when they happened. However, because there is no goal setting for incident report, the system does not support the type of self-monitoring that allows the user to track their progress after setting a report-based goal or when submitting a report, for example, through the display of a progress bar. In addition, the system does not allow the user to track the levels of usefulness of their reports (eg, incident, near miss, and CQI plans) to others. For example, it does not allow other users or colleagues to "like" the user's anonymous reports or to indicate their usefulness.

Tailoring, Personalization, and Customization

All 3 persuasive strategies are related and can be defined as the act of tailoring the user interface elements and content of a system to suit the user's needs, preferences, designation, or role. Tailoring and personalization are carried out by the system, whereas customization is carried out by the user. Although tailoring is enacted by the system based on users' predetermined information (eg, gathered through surveys before using the system), personalization is enacted by the system using information gathered in real time (ie, during user interaction with the system) [81,83]. We observed that tailoring was implemented in the assessed system. This system provides

XSL•FO
RenderX

role-based access to certain features. However, the assessors remarked that the tailoring feature can be improved depending on what users need. However, we found that the system does not support personalization. Hence, we recommend MIRLS be personalized based on user interaction, for example, letting the user know where they left off or reminding them about incomplete tasks when they log in [7,66]. In addition, we recommend that users be allowed to customize the system (eg, user profile, chart, content, information, and reminder) to suit their needs and preferences [66].

Simulation

Simulation is a persuasive strategy used to demonstrate the cause and effect of a given behavior. Although the assessors considered it important, it was not currently implemented in the system. Thus, we recommend that MIRLS provide a means for the user to observe a link between the cause and effect of incident and near-miss reporting [80,84]. A typical implementation of the strategy is demonstrating to the user using a graph or chart that the higher the MIs reported using the system, the lower the number of recurrences.

Rehearsal

Rehearsal is a trial performance or practice of a given task so that the user can perform it correctly and easily later. In the assessment, we found that the system already provides a new user with video tutorials (organized in modules) on how to report a MI or near miss. In addition, we recommend that MIRLS provide a new user with a practice environment, in which they can rehearse before using the system to make an actual report.

Dialogue Support Assessment and Guidelines

In the dialogue support category (Table 3), all the persuasive strategies were considered important or useful; however, only 50% (4/8) of them (eg, reminder, feedback, and suggestion) were considered partially implemented by the system.

Praise and Reward

They entail acknowledging, appreciating, and recognizing the user for their effort and time taken to report incidents and near misses for the benefits of other pharmacists and patient safety. As Holden et al [6] noted, “reward and punishment structures may affect individual reporting decisions (e.g. if nurses are rewarded more for productivity than for reporting), as may culture (e.g. blame vs. just culture).” It is yet to be seen how the web-based rewards implemented in a system may influence error reporting. Enacted through well-worded motivational text and well-designed motivational images, symbols, and sounds [67], praise fosters an intimate relationship between the user and the system, making the user feel valued, appreciated, and more open to persuasion [85]. Although considered important by 2 of the assessors, 1 of them had some reservation. The participant stated, “This would emphatically not be wanted. Reward messages coming from an MIRLS technology should not emulate a sports watch. As an advanced user of the system I would find this annoying and a waste of time. If the system helps reduce incidents, a trend report shows proof, that is praise enough.” However, praise and rewards can be targeted to aggregated reports (eg, a pharmacy) on the basis of the number of incidents that reached and did not reach the patient.

Suggestion

This strategy is considered important and partially implemented in the system and can be used as a means of informing users about certain important reports (especially from other anonymous users or generated from the system), which may be useful to them in their practice. A typical suggestion in this context could be a list of “actions to take” for a specific MI or a list of “high-alert” medications that require extra precautions. Other suggestions include new research reports that may be interesting and beneficial to the user or ways, processes, or methods through which other anonymous users in the community prevent or address recurrence of certain medication errors [65]. For example, upon completing a report, the user can be recommended a set of preventive guidelines by the system to mitigate future incidents.

Feedback

Several behavior change theories such as social cognitive theory, goal setting theory [86], and feedback intervention theory consider the provision of feedback as an important ingredient in behavior change [87,88]. An example implementation of the self-monitoring-type of feedback is providing the user with summary feedback on their progress toward reaching their goal (eg, “You have achieved 30% of your goal”). Moreover, feedback entails information about one’s behavior or system-generated figures and statistics. In the context of MIRLS, informational feedback is the information of the user about the impact of their error-reporting behavior on the community or health providers’ medication errors on patient safety. An example of informational feedback is informing pharmacists about the usefulness of their reports to other users in the community (eg, “5% of the system users in the province [nation] found your report helpful”). Another example is providing users with monthly, quarterly, or yearly summary feedback highlighting the most recurring types of errors relevant to their work [65] (eg, “Poor drug naming caused 5% of the near misses last year.” In addition, the solution to this medication error can be included in the feedback message as well; for example, “Poor drug naming caused 5% of the near misses last year; remember to use TALLman lettering when necessary.” The use of uppercase letters in a portion of a drug name helps to draw attention to the dissimilarities between look-alike and sound-alike drug names. Moreover, it helps to alert health care professionals that the name of a given drug can be confused with another drug that has a similar name [89].

Reminder

This refers to an alert on task completion and compliance with certain behavior or expectation [90]. Reminder is closely tied to goal setting in a certain regard. For example, if the user sets a goal (eg, report at least X errors per month), then the user should have the opportunity to set reminders so that they could be reminded at certain preset times to report incidents or near misses if they have any. Reminder has been widely and successfully used in persuasive systems, especially in the health domain, to motivate behavior change [80,91]. In MIRLS, reminders, considered important and partially implemented,

XSL•FO
RenderX

can be based on users' self-set goals on medication error reporting as well as CQI-based action plans. For example, based on self-set goals, the system can remind the user at preset times about the need to report near misses and incidents when they occur and about the benefits of the reports to other users in the community and patient safety. For instance, the system can prompt the user at a preset time with a message such as, "Did you have any near misses today or in the last one week? Please report if you did." Moreover, the system can remind the user through this type of message if the user has not logged into it or submitted a report within a certain period. In addition to this reminder-based messages, a direct link to a reporting wizard can be included, allowing users to easily submit a report by simply clicking on the provided link. Persuasive reminders have been widely used in health self-management such as taking one's daily medication and have been effective [92]. Although reminders may be more effective if they are just-in-time [87], in the context of MIRLS, they can be well ahead of time, for example, during the period when a user such as a pharmacist resumes their shift. They can also be at the end of the pharmacist's shift. Therefore, research, in the context of MIRLS, is required to show which of the periods (start or end) is more likely to be effective in motivating reporting of medication errors. In summary, reminders can be general or specific. General reminders are aimed to remind users from time to time to report incidents if they have any. Moreover, specific reminders are aimed to remind users to complete incident report drafts (ie, reports that they started but have not completed). Nevertheless, reminders should be used with caution as they can be overwhelming if overused. As stated by 1 of the assessors, "Reminders can also be annoying to the point of reminder fatigue and disregarded instantly, and overkill for this type of solution." Therefore, users should be allowed to turn them on and off.

Verbal Persuasion

This refers to the act of mentoring and providing encouragement and feedback to help individuals achieve their goals. It is also defined as "the act of telling or convincing a person to perform a task or action to change a behavior or put into action a set of events to achieve an objective" [93]. Research shows that organizational and leadership coaches use verbal persuasion effectively to increase the self-efficacy of their clients and the results they create. The tools for carrying out verbal persuasion include praise (kind words about the user), encouragement (words of affirmation about the user's ability), stories (personal or allegorical stories to help reframe the user's struggle with the task), positive feedback (assessing the user's performance favorably), strengths focus (intentionally linking the task to the user's strengths), and past achievements (acknowledging past wins as an indication of the user's ability to complete the current task) [94]. In the context of MIRLS, praise and encouragement may be used effectively by community pharmacy managers and supervisors to motivate users to report near misses and incidents. However, the use of individual feedback and past achievements may not be possible in MIRLS if, at the pharmacy level, managers and supervisors do not have access to individual users' performance owing to anonymity. In the event that managers had access to individual users' performance, as may be the case in certain pharmacies owing to corporate policy, managers and supervisors could enact verbal persuasion through personal feedback and strengths in addition to praise and encouragement. Although verbal persuasion can be said to be related to the praise and emotional appeal strategies, the main difference is that verbal persuasion is coming directly from a superior (eg, a pharmacy manager) that the user knows rather than the system. A typical message a pharmacy manager can send to an employee to verbally persuade them is, "Alice, remember to report your near misses and incidents to improve patient safety. Yes, you can!" Moreover, a typical feedback message from a pharmacy manager is, "Alice, thanks for your constant reporting of near misses—keep it up!" Users (whether reporting frequently or not) may find this type of message motivational. This may motivate users who have not been reporting their errors using the system in recent times to start reporting. Moreover, this type of positive feedback will help address one of the administrative barriers presented in Textbox 1: "Underreporting due to lack of useful feedback or negative feedback from administrative teams such as pharmacy managers" [8].

Emotional Appeal

It is a persuasive strategy designed to elicit an emotional response based on feelings [95]. We argue that motivational messages that appeal to emotion and feeling, such as "To err is human, to share is divine" [69], have the potential to motivate users in the medication error–reporting domain, similar to other domains [81]. In the fitness app domain, for example, Oyibo [96] found that, regardless of gender, health messages that appeal to emotion, such as "Those who do not find time for exercise will have to find time for illness," have the potential to motivate people to start or continue exercising. However, in this study, we found that although a motivational message such as "To err is human, to share is divine" may motivate some pharmacists, as evident in 1 of the assessors' responses ("would love it"), it may demotivate others. One of the assessors commented that the use of emotional appeal is inappropriate in a professional domain such as community pharmacy. The assessor stated, "It is a regulatory requirement to report incidents–no need for motivational messages...like a sports watch or fitbit. It seems unprofessional for a tool such as this to have this. I would NEVER accept this or turn this feature on." The mixed reactions to the use of emotional appeal to motivate incident reporting, similar to praise and reward, require further empirical studies.

Liking

This entails making a system visually attractive and engaging to make it persuasive. This strategy in the PSD taxonomy is drawn from the 6 principles of persuasion proposed by Cialdini [73]. According to Cialdini [73], the more people like someone, the more likely they are to be persuaded by the person. Similarly, in the context of PSD, the more esthetic a system is, the more persuasive users find it and the more likely the users are willing to use it to motivate their behavior change [48,97]. In the context of MIRLS, designers can use visually pleasing user interfaces and appropriate colors to present charts, content, and important information to improve the overall UX.

XSL•FO
RenderX

Social Support Assessment and Guidelines

Overview

In the social support category (Table 3), we only found that social comparison (in the form of benchmarking) was already implemented in the system for a limited number of measurements. However, the user had to filter each time to be able to benchmark the measure of interest (eg, near miss) at one level (eg, in the pharmacy) against another (eg, in the province). The assessors of the system suggested that rather than filtering all the time, it would be better if the benchmarking feature of the system could be enhanced by locking in the error reports—having them appear automatically. Moreover, we recommend guidelines on how to integrate other socially oriented persuasive strategies such as social learning, social facilitation, normative influence, competition, and social recognition. Holden and Karsh [7] found that social influence at the individual, group, organizational, and industry levels has the potential to influence medication error reporting.

Social Learning

This social strategy allows users to observe and imitate the behaviors and achievements of other (anonymous) users of the system [98]. The social learning strategy derives from the social learning theory proposed by Bandura [99]. The social learning theory states that people have the ability to imitate new behavior by coding or storing the ideas about the behavior in their memory, which eventually guide the actual performance of the behavior [100]. In the context of persuasive technology, social learning is simply implemented using the information of the target user about a target behavior performed by other users, for example, through a notification. In the context of MIRLS, a potential approach to implementing social learning is by enabling users to receive notifications (eg, via email) when fellow users in their group submit incident reports. These notifications would contain essential key points from the submitted reports. A typical notification message to this effect is “John [a pseudonym], here’s a new report we think you’ll be interested in.” We believe that messages such as this, which enable one user to learn from others’ reports, may motivate the target user to submit their reports given the benefit they derive from them. Given that users may be overwhelmed, they should be given the opportunity to determine the types of messages they wish to receive, the number within a given period such as a week or month, and even opt out completely by turning the feature off. More importantly, owing to privacy concerns, particularly within a facility setting, instead of basing the social learning strategy on key points from reported near misses or incidents, it can be based on the quantity of reports submitted within a specified period (refer to the *Normative Influence* section). According to 1 of the assessors, “I don’t think this [first Social Learning implementation] is appropriate if you can see who it is but if it is just numbers it would be useful. [N]otification within a facility could hamper the feeling of safe reporting because anonymity is compromised.” A second implementation of social learning is the provision of a news feed that highlights important reports submitted by other anonymous users that the user may find useful. A third implementation is the support of chat rooms or discussion rooms where users can discuss near misses, incidents and lessons learned; share experiences and knowledge; and learn from one another in an anonymous fashion. The chat room and discussion forum feature may be extended and beneficial to nonpharmacists, as evident in 1 of the assessors’ comments, “Our users may find this useful. If they have the time, which currently they don’t have much of during the pandemic.”

Social Comparison

Social comparison allows users to compare their performance with that of others. It is derived from the social comparison theory proposed by Festinger [101], which centers on the belief that individuals have an inner drive to gain accurate self-evaluations through social comparison. It holds that by comparing one’s abilities and performances with those of similar others or peers, the individual is able to reduce uncertainty, learn, and improve self. This strategy has been used successfully in persuasive systems [102]. In the assessment of the Think Research or Pharmapod system, we found that social comparison was implemented at the pharmacy and provincial level in the form of benchmarking reports, tables, and dashboards. For example, 1 of the assessors responded thus, “within our own organization we may compare pharmacies with other pharmacies or between provinces of our pharmacies using reports provided.” Thus, the implementation of social comparison in the system can be improved. For example, users’ error reporting over a particular period, for example, week, month, or year, can be compared anonymously with the average at the pharmacy or provincial level using a bullet chart infographic.

Competition

Similar to social comparison, competition allows users to compare themselves with others, for example, in terms of number of reports, frequency, quality, or usefulness of reports to others. Competition leverages the natural drive of humans to outperform one another [98]. Research on persuasive technology shows that competition, regardless of gender, age, and culture, has the potential to motivate users to perform the target behavior [103]. In the fitness app domain, for example, Oyibo and Vassileva [98] found a significant relationship among social comparison, social learning, and competition, indicating that the more people compare themselves, the more they learn about the performance or achievements of others and the more competitive they become in their behaviors. In the context of MIRLS, users can be allowed to view where they are compared with other anonymous users in small sets (eg, on a leaderboard). The criterion for placement on the leaderboard can include the total number of reports, frequency, quality, or usefulness of the report to others (eg, over a weekly, monthly, or yearly period). The small sets of anonymous users can be drawn from the pool of users at the provincial or national level, which can change from time to time because of the need to foster anonymity. Moreover, the competition feature can be group based, involving anonymous pharmacies, organizations, or provinces. As 1 of the assessors remarked, “Perhaps [my organization] may wish to see how many incidents they are experiencing compared to another organization of the same industry channel and size.”

XSL•FO
RenderX

Cooperation

Unlike competition, where users compete to outperform one another, in cooperation, users work together in a collaborative fashion to achieve their individual and collective goals. In the assessment of the Think Research or Pharmapod IM system, we found that providing users the choice of being paired with another (anonymous) user, with the goal of motivating one another to achieve individual or collective goals may not be a good idea. This is based on the premise that the implementation of cooperation in MIRLS may compromise the principle of anonymity of users, upon which MI reporting is founded. Hence, we recommend that cooperation be implemented and used with caution if MIRLS were to support it in a given pharmacy. As commented by 1 of the assessors, "Why would anyone wish to be compared to [cooperate with] another user? Where's the privacy aspect of such a feature?"

Normative Influence

Unlike informational influence, which is conformity to a certain behavior based on the acceptance of evidence about reality provided by others, normative influence is conformity based on an individual's desire to fulfill others' expectations to gain acceptance, fit in, or feel a sense of belonging [104]. In the context of reporting medication errors, the urge for individual users to report near misses and incidents might arise from perceived social pressure rather than actual pressure, considering that the submitted reports are anonymous or deidentified. Thus, a possible way of realizing the normative influence strategy in MIRLS is allowing the user to know about the number of other anonymous users in the facility, province, or nation that are reporting medication errors at a given time. For example, in COVID-19 contact tracing apps, Oyibo and Morita [105] found that socially oriented messages, such as "112 other people reported their COVID-19 diagnosis today," have the potential to motivate app users to report their diagnosis by entering their one-time key into the app. Hence, we recommend that the system informs users at suitable intervals (eg, when they are logged on) about the quantity of other anonymous users within the pharmacy, province, or country who are reporting medication errors within a specific period. A message similar to the message by Oyibo and Morita [105], "10 other people submitted their incident reports today," may be used to normatively influence users to submit their own incident reports as well if they have any pending or have not yet submitted.

Social Facilitation

Social facilitation refers to the improvement in a person's performance as a result of the real, imagined, or implied presence of others. As stated in the study by Mohadis et al [84], "System users are more likely to perform a targeted [behavior] if they discern, via the system, that others are performing the [behavior] along with them." In MIRLS, one way to realize social facilitation is to inform the user when they log on to the system (eg, to make a report) through news feed that they are not alone in their efforts to report an error, as other users elsewhere (eg, in the facility, province, or nation) at the current time are also attempting to making a report or logged on to the system. Motivational messages such as "You are not alone; X others are on the system at the moment submitting a report" could be used to make the user feel the presence of other anonymous users whenever the former is logged into the system. A message such as this may encourage users, who have begun the process of submitting a report, to complete it. This type of message is similar to that which customers get when they are booking a hotel or shopping for a flight ticket on the web (eg, "5 other people are currently shopping for this flight ticket"). Although this type of message is commonly used in the e-commerce domain to create the impression that the user may miss procuring a given flight ticket if they do not act quickly (ie, buy it now), in the domain of medication error reporting, this is not the case. Rather, this type of message is used to let the user know that they are not alone—that there are similar others elsewhere who are trying to do the same task as them (submit a MI report).

Social Recognition

In social psychology, social recognition is the act of recognizing people such as employees for great work, contribution, and achievement by acknowledging them publicly. One possible way of implementing this strategy in an MIRLS is recognizing users for being one of the "most valuable players" of the month, quarter, or year. This can be at the facility, provincial, or national level. The criteria for recognition include the number, frequency, quality, or usefulness of the target user's reports to the community. Although research shows that employees welcome social recognition in the workplace [84], it must be implemented with caution given the anonymity requirement aimed to protect users from punitive measures. We found that users may not welcome the second feature ("allowing other users to rate a user's report anonymously based on how useful or helpful it is to them") as they perceived it as a form of competition. For example, 1 of the assessors commented, "Rating makes this feel like a competition or to call out that can produce negative attitudes. Not helpful. Those entering data into a system may not be the same person who is involved in the incident." Moreover, the user was also concerned about the part of the report being rated as well as privacy and anonymity, "What part of the report is being rated in this scenario?" It is worth noting that we conceived the social recognition rating feature similar to Google Play Store app rating system, in which users can rate an app on a 5-star scale. Although we did not explicitly detail the section of the report being anonymously rated by other users in the study, we intended it to encompass essential elements derived from the report analysis, such as the description of the near miss or incident, the lessons learned by the reporters, and possible recommendations and tips to prevent future recurrence. These key points may have been extracted from a set of similar aggregated reports submitted by different anonymous users at different times and included in the MI analysis report shared with users via the MIRLS by standard bodies such as Assurance and Improvement in Medication Safety (AIMS) [106]. AIMS is a standardized medication safety program that supports CQIs and sets a mandatory consistent standard for medication safety for all pharmacies in Ontario. Its goal is to minimize the risk of harm to patients caused by MIs in the province. Part of its mandate is to aggregate and analyze anonymous MI reports and produce and disseminate the results to stakeholders. This enables practitioners to learn from MIs and have a better understanding

XSL•FO
RenderX

of why they occur and how they can be prevented in the future [106]. Although in this study, we did not find the second social recognition feature to be useful to the assessors, there may be a need for a more comprehensive study in future research among a larger audience of community pharmacists to uncover its potential to motivate users to report medication errors more frequently.

System Credibility Support Assessment and Guidelines

Regarding the credibility support category (Table 4), the assessors reported that the system fully or partially supported a number of credibility-related persuasive strategies such as trustworthiness, credibility, expertise, and real-world feel. We discuss all these strategies together with the other 3 strategies in the credibility support category.

Authority

One of the principles of persuasion proposed by Cialdini [73], the authority principle, states that people are more likely to believe and obey those who are in positions of authority. Selassie et al [76] found that frontline staff working with children with autism (supported by a data entry management system) can be persuaded by the authority strategy. Moreover, in the study by Mohadis and Ali [84] on user perception of a physical activity app for older workers, 1 of the participants remarked, "Yeah, incorporating an expert [authority figure's] view is very important so that we become more confident with whatever recommendations that the system offers." In the context of community pharmacy, authority figures and bodies may include researchers, pharmacy managers, and professional bodies such as the Institute for Safe Medication Practices Canada [1]. Thus, we recommend the presentation of authority-based information and messages to users, for example, on the value of reporting medication errors and the benefits it can have for the profession, staff, and patient safety [47].

Third-Party Endorsement

Third-party endorsement is the act of publicly approving or supporting a product, system, or service by a reputable socially influential individual or organization other than the staff or company that owns it. Usually, the third party may have seen, interacted, and used the product, system, or service in question and is satisfied with the results, utility, or experience. In the business world, research has shown that the third-party endorsements have the potential to effectively earn companies the trust and loyalty of customers [21,69]. Moreover, research shows that the expertise and trustworthiness of a third-party organization endorsement have the potential to positively affect the perceived value of a firm, which in turn can positively affect customer loyalty [107]. Hence, to encourage pharmacists to use MIRLS, the designers should demonstrate that the system is approved or endorsed by authoritative bodies such as professional organizations (eg, World Health Organization and Institute for Safe Medication Practices), regulatory bodies, and government. To implement this persuasive strategy in MIRLS, one approach is to incorporate the corporate logos of the endorsing authoritative bodies within the user interface, such as on the system's home page or in the footer, especially if it is a web-based application.

Expertise, Surface Credibility, and Trustworthiness

Research has shown that all 3 strategies are related. For example, Fogg and Tseng [108] postulated that credibility, a perceived quality of a system, comprises 2 key components: trustworthiness and expertise. In other words, a system is perceived to be credible if its perceived trustworthiness and perceived expertise are high. Trustworthiness is a key element in the credibility perception of systems such as websites. It is defined by terms such as well intentioned, truthful, and unbiased [109]. As stated in the study by Fogg et al [109], "the trustworthiness dimension of credibility captures the perceived goodness or morality of the source." Similarly, expertise is a key element in the credibility perception of systems such as websites. It is defined by terms such as knowledgeable, experienced, competent, and professional [109]. As stated in the study by Fogg et al [109], "[t]he expertise dimension of credibility captures the perceived knowledge and skill of the source." In a large-scale website credibility study conducted by Fogg et al [109], the authors found that perceived expertise and perceived trustworthiness have a significant impact on the perceived credibility of websites. In the context of MIRLS, to realize expertise, the visual and functional design of the system should reflect professionalism, expertise, and up-to-dateness to motivate users to use it. Moreover, to implement trustworthiness, the system should foster user anonymity, data deidentification, and data aggregation and live up to promises such as it not being used as a punitive tool to hold users accountable [65]. Finally, perceived credibility can be intentionally built into the system through its visual design, for example, by ensuring users enter accurate information using taxonomy-based option buttons, checklists, and drop-downs and reducing advertisements for a web-based system [47]. In our study, all 3 assessors agreed that perceived expertise is important or useful as well as implemented to a great extent in the system they were currently using. For example, 1 of the assessors commented, "The MIRLS is very easy to use and intuitive, and requires minimal training to get started." However, "there is always room for improvement," remarked another assessor. Failure to foster expertise in the system design may discourage frequent use and completion of tasks, as evident in the assessor's comment, "Performance in speed is always a challenge and [the] latency [experienced in some] areas drive users to drop off or stop using." Regarding trustworthiness, 2 assessors considered it important or useful. However, only 1 assessor considered it to be implemented in the current system. This is partly because of anonymity not being completely fostered in the system. This is evident in 1 of the assessors' comments, "anonymity is fostered outside an organization (eg. when data sent to AIMS) and there is also a choice to report anonymously so the corporate level of an organization does not have visibility. [W]ithin a location the reports are not anonymous." Finally, regarding surface credibility, 2 assessors considered it important or useful and implemented it in their current system. For example, all 3 assessors responded that there were no advertisements in the system and that was very important.

Verifiability

This refers to "the quality or state of being capable of being verified, confirmed, or substantiated" [110]. In the context of

XSL•FO
RenderX

MIRLS, persuasive messages (eg, on the value of error reporting to patient safety) aimed at motivating users should not only be credible but also verifiable. As stated in the study by Jones [111], carefully choosing persuasive messages and supporting materials that are verifiable, specific, and unbiased can be helpful in appealing to logic and increasing users' trust. Verifiability was implemented in WargaFit (a fitness app prototype aimed to encourage simple exercise such as body stretching in an office environment) by the provision of healthy tips accompanied with external links [84]. Similarly, verifiability in MIRLS can be realized through the provision of the source of information or inclusion of the URL in the persuasive message such as "Reporting reduces the number of future errors, diminishing personal suffering and decreasing financial costs" [112]. In our study, 2 assessors considered verifiability useful and not currently implemented in their system. For example, regarding harm levels, 1 of the assessors commented, "There are info points that explain [that] harm level comes from WHO but there is no link to the WHO to verify it."

Real-World Feel

Similar to expertise and trustworthiness, real-world feel is found to positively influence the perceived credibility of websites [109]. Real-world feel is the interaction with and experience of a virtual or electronic product, system, or service as though it is real. This is made possible by the product-, system-, or service-supporting features that mimic and foster real-world interaction and experience. In the case of e-commerce websites, for example, the real-world feel can be fostered by providing contact phone number, contact email address, and a quick response to customer service questions; listing the physical address of the organization behind the website; and showing photos of the members of the organization [109]. In the context of MIRLS, in addition to the aforementioned features, the system should be designed as close as possible to the nonelectronic (paper) version. This has the potential to reduce the cognitive effort required by a new user to make the transition. In the assessment of the Think Research or Pharmapod IM system, assessors stated that it supports real-world feel by mimicking the paper version and allows clients to customize their own forms and notifications or escalations. One way the system designers achieved real-world feel is to allow pharmacies and organizations to customize their MI report forms.

Persuasive System Implementation and Ethical Design Considerations

Our analysis reveals that there is a need to consider and address the ethical implications that may arise from integrating persuasive strategies into the existing MIRLS. These considerations include administrative (eg, anonymity) and choice of persuasive strategies (eg, monetary reward). For example, to ensure that the principle of anonymity is fostered in the implementation of social strategies, user identifications should be limited to pseudonyms, which the users can change from time to time. It is worth noting that a persuasive strategy that may be effective (or welcomed) in one community pharmacy may not be in another. Hence, there may be a need to get the potential users involved in deciding the set of persuasive strategies that will be implemented or effective in a given pharmacy. Thus, the system should offer tailoring capabilities that support the chosen guidelines. Intervention researchers and designers may have to (1) investigate, before implementation, which of the recommended persuasive strategies a given group of pharmacy professionals may be or may not be receptive to and (2) implement only the set of strategies that are likely to be effective, as proven by empirical evidence. For these reasons, MIRLS should be designed in a way that enables pharmacies to turn on and off persuasive strategies that they consider useful and nonuseful, respectively. It is worth noting that some of the persuasive strategies in the PSD taxonomy may have to be combined to realize a holistic and functional persuasive feature that is useful. In other words, some of the persuasive strategies are complementary. For example, praise and feedback strategies must be combined to implement or realize a composite feature that provides immediate feedback of praise to the user upon submitting an incident report. In addition, reminders and verbal persuasion may be combined to realize a composite feature that verbally persuades the user through a reminder. For example, a verbal persuasion message ("Alice, remember to report your near misses and incidents to improve patient safety. Yes, you can!") can be sent to the user as a reminder by the pharmacy manager from time to time. Finally, authority, credibility, and verifiability may have to be combined to realize a persuasive message that is not only authoritative and credible but also verifiable.

Contributions

In this study, we have made a number of contributions to knowledge in the domain of community pharmacy and developers of health digital systems. This study is the first to provide guidelines on how to integrate persuasive strategies into MIRLS to increase their utility and motivate users to report MIs and near misses to improve patient safety and promote shared learning. Specifically, we provided MIRLS-specific persuasive design guidelines based on the PSD taxonomy proposed by Oinas-Kukkonen and Harjumaa [30]. Most of the PSD guidelines in the extant literature are concentrated in the domains of healthy eating [113] and physical activity [81,84,114]. Designers of MIRLS can leverage the current set of PSD guidelines in improving future iterations not only in community pharmacy but also in other settings where incident or error reporting is essential and part of the organizational practice. The second contribution is that this study lays the foundation for future empirical research aimed at investigating the effectiveness of persuasive strategies incorporated into MIRLS. Future research efforts should focus on ≥1 of the design guidelines in each of the 4 categories of the PSD taxonomy; implement them; and conduct a field study to examine the perception, acceptance, and adoption of the implemented strategies by the target community pharmacists.

Research Directions

In future work, we look forward to investigating the potential effectiveness of some of the proposed persuasive design guidelines presented in Tables 2-4 and Textbox 4 in field studies. First, we will create prototypes of the persuasive strategies and perform an empirical study to explore which set of strategies might be more effective. In addition, we will analyze the

XSL•FO
RenderX

potential influence of demographic variables, such as age, gender, and work experience, on the effectiveness of these strategies. Second, we will select the most persuasive strategies that the target community of pharmacy professionals are most responsive to and implement them in an actual MIRLS (eg, Think Research or Pharmapod). Third, we will conduct a field study (randomized controlled trial) to investigate the effect of the persuasive design on the rate of MI reporting among community pharmacy professionals using different provinces across Canada as case studies. More importantly, owing to the lack of studies on the relationship between system usability and medication error reporting, as our second scoping review shows, we recommend that future work be conducted in this area.

Limitations

Similar to most conceptual papers, our study has limitations owing to its preliminary nature, which stems from the nonmaturity of research on the persuasive design of MIRLS. The first limitation is that the results of the scoping reviews might have been limited one way or the other by the choice of search strings and the subjective assessment, understanding, and interpretations of the extracted data by the researchers that conducted the reviews. Hence, we recommend a more comprehensive review, particularly with regard to the second RQ, in which a formal review led to no included article, other than the article retrieved from Google Scholar search. The second limitation of our study is the convenience sample. In other words, the 3 assessors who assessed the Think Research or Pharmapod system using the PSD taxonomy were not sufficient to be representative of the entire population of community pharmacy professionals using MIRLS across Canada. For example, a persuasive feature that may be important and useful to a group of community pharmacists in one facility may not be useful to another group in another facility. Hence, the findings reported in the last 2 columns of Tables 2-4 and Textbox 4 may not generalize to a larger population sample involving a heterogeneous group of community pharmacists with different roles, working environments, years of working experience, professional qualifications, gender, personality, and economic status, which may influence their responses. In future work, we hope to build on this preliminary study by conducting a formal research (eg, based on storyboards) involving a larger population sample to validate the generalizability of the findings of this study, particularly the effectiveness, acceptability, and adoption of the recommended persuasive strategies presented in Tables 2-4 and Textbox 4.

Conclusions

Although most medical practitioners agree that reporting medication errors improves the quality of care and safety for patients [21], in reality, the rate of reporting remains below expectations [115] owing to lack of motivation and other barriers [22-25]. In this study, we argued that although most current MIRLS have implemented recommended guidelines bordering on favorable administrative measures and utility, they lack motivational affordances that can facilitate or motivate frequent reporting. Hence, using the Think Research or Pharmapod system as a case study, we identified opportunities for incorporating persuasive strategies into MIRLS to make them more effective in motivating behavior change. The proposed persuasive design guidelines can be used by designers and developers in making MIRLS more effective in motivating users to report incidents and near misses more often to reduce risks of recurrence, improve patient safety, and foster shared learning among community pharmacy professionals and stakeholders. However, before the implementation of the recommended persuasive design guidelines in Tables 2-4 and Textbox 4, there is a need for thorough consideration and evaluation of the various ramifications, including administrative, regulatory, and ethical implications. The presented persuasive design guidelines open up new opportunities for persuasive design research in MI reporting. We acknowledge that some of the proposed persuasive strategies may not be suitable or effective in real-life settings. Hence, there is a need for further validation-based research and caution regarding their implementation. In future work, we aim to validate the suitability and effectiveness of the proposed persuasive strategies in motivating behavior change using storyboards, prototypes, and perception and evaluation studies involving community pharmacists across Canada.

Acknowledgments

This project was funded by the Social Sciences and Humanities Research Council Partnership Engage Grant (#1008-2020-0119).

Authors' Contributions

KO and PAG coordinated and supervised the assessment of the Think Research or Pharmapod platform and led the substantive writing of the paper. KO, SE, and TN conducted the scoping reviews including database search, retrieval of articles, screening and selection of included articles, and extraction and tabulation of data. CB facilitated access to the platform, contact with the assessors, and reviewed and edited the paper. DO provided technical assistance in the platform access and assessment and reviewed and edited the paper. KO used the data from the scoping reviews and assessment of the Think Research or Pharmapod platform to write the paper. PAG and JRB contributed to the editing of the paper.

Conflicts of Interest

CB is the vice president of Quality Improvement and Innovations at Think Research or Pharmapod, which is the partner organization for the Partnership Engage Grant that funded this research. All other authors declare no other conflicts of interest.

References

XSL•FO
RenderX

1. Deaths associated with medication incidents: learning from collaborative work with provincial offices of the chief coroner and chief medical examiner. Institute for Safe Medication Practices Canada. 2013 Aug 28. URL: https://tinyurl.com/fycmya97 [accessed 2024-01-11]
2. Alqenae FA, Steinke D, Keers RN. Prevalence and nature of medication errors and medication-related harm following discharge from hospital to community settings: a systematic review. Drug Saf 2020 Jun;43(6):517-537 [FREE Full text] [doi: 10.1007/s40264-020-00918-3] [Medline: 32125666]
3. Desjardins L. Thousands die from medical errors yearly, notes advocacy group. Radio Canada International. 2019 Oct 28. URL: https://tinyurl.com/bdcepsmu [accessed 2023-08-20]
4. Tariq RA, Vashisht R, Sinha A, Scherbak Y. Medication dispensing errors and prevention. In: StatPearls Internet. Treasure Island, FL: StatPearls Publishing; 2013.
5. Mrayyan MT, Shishani K, Al-Faouri I. Rate, causes and reporting of medication errors in Jordan: nurses' perspectives. J Nurs Manag 2007 Sep;15(6):659-670. [doi: 10.1111/j.1365-2834.2007.00724.x] [Medline: 17688572]
6. Kim KS, Kwon SH, Kim JA, Cho S. Nurses' perceptions of medication errors and their contributing factors in South Korea. J Nurs Manag 2011 Apr;19(3):346-353. [doi: 10.1111/j.1365-2834.2011.01249.x] [Medline: 21507105]
7. Holden RJ, Karsh BT. A review of medical error reporting system design considerations and a proposed cross-level systems research framework. Hum Factors 2007 Apr;49(2):257-276. [doi: 10.1518/001872007X312487] [Medline: 17447667]
8. Aljabari S, Kadhim Z. Common barriers to reporting medical errors. ScientificWorldJournal 2021 Jun 10;2021:6494889 [FREE Full text] [doi: 10.1155/2021/6494889] [Medline: 34220366]
9. Bahadori M, Ravangard R, Aghili A, Sadeghifar J, Gharsi Manshadi M, Smaeilnejad J. The factors affecting the refusal of reporting on medication errors from the nurses' viewpoints: a case study in a hospital in Iran. ISRN Nurs 2013 Apr 09;2013:876563 [FREE Full text] [doi: 10.1155/2013/876563] [Medline: 23691354]
10. Yung HP, Yu S, Chu C, Hou IC, Tang FI. Nurses' attitudes and perceived barriers to the reporting of medication administration errors. J Nurs Manag 2016 Jul;24(5):580-588. [doi: 10.1111/jonm.12360] [Medline: 26888342]
11. Higuchi A, Higami Y, Takahama M, Yamakawa M, Makimoto K. Potential underreporting of medication errors in a psychiatric general hospital in Japan. Int J Nurs Pract 2015 May;21 Suppl 2:2-8. [doi: 10.1111/ijn.12169] [Medline: 26125569]
12. Alqenae FA, Steinke D, Carson-Stevens A, Keers RN. Analysis of the nature and contributory factors of medication safety incidents following hospital discharge using national reporting and learning system (NRLS) data from England and Wales: a multi-method study. Ther Adv Drug Saf 2023 Mar 16;14:20420986231154365 [FREE Full text] [doi: 10.1177/20420986231154365] [Medline: 36949766]
13. George D, Hassali MA, Hss AS. Usability testing of a mobile app to report medication errors anonymously: mixed-methods approach. JMIR Hum Factors 2018 Dec 21;5(4):e12232 [FREE Full text] [doi: 10.2196/12232] [Medline: 30578216]
14. David U. Medication error reporting systems: problems and solutions. New Med 2001 Nov;1(2):61-65.
15. Kang HJ, Park H, Oh JM, Lee EK. Perception of reporting medication errors including near-misses among Korean hospital pharmacists. Medicine (Baltimore) 2017 Sep;96(39):e7795 [FREE Full text] [doi: 10.1097/MD.0000000000007795] [Medline: 28953611]
16. West DR, Pace WD, Dickinson LM, Harris DM, Main DS, Westfall JM, et al. Relationship between patient harm and reported medical errors in primary care: a report from the ASIPS collaborative. In: Henriksen K, Battles JB, Keyes MA, Grady ML, editors. Advances in Patient Safety: New Directions and Alternative Approaches (Vol. 1: Assessment). Rockville, MD: Agency for Healthcare Research and Quality; Aug 2008.
17. Robb G, Loe E, Maharaj A, Hamblin R, Seddon ME. Medication-related patient harm in New Zealand hospitals. N Z Med J 2017 Aug 11;130(1460):21-32. [Medline: 28796769]
18. Riordan CO, Delaney T, Grimes T. Exploring discharge prescribing errors and their propagation post-discharge: an observational study. Int J Clin Pharm 2016 Oct;38(5):1172-1181. [doi: 10.1007/s11096-016-0349-7] [Medline: 27473712]
19. Panagioti M, Khan K, Keers RN, Abuzour A, Phipps D, Kontopantelis E, et al. Prevalence, severity, and nature of preventable patient harm across medical care settings: systematic review and meta-analysis. BMJ 2019 Jul 17;366:l4185 [FREE Full text] [doi: 10.1136/bmj.l4185] [Medline: 31315828]
20. Boyle TA, Mahaffey T, Mackinnon NJ, Deal H, Hallstrom LK, Morgan H. Determinants of medication incident reporting, recovery, and learning in community pharmacies: a conceptual model. Res Social Adm Pharm 2011 Mar;7(1):93-107. [doi: 10.1016/j.sapharm.2009.12.001] [Medline: 21397884]
21. Vrbnjak D, Denieffe S, O'Gorman C, Pajnkihar M. Barriers to reporting medication errors and near misses among nurses: a systematic review. Int J Nurs Stud 2016 Nov;63:162-178. [doi: 10.1016/j.ijnurstu.2016.08.019] [Medline: 27637011]
22. Elder NC, Graham D, Brandt E, Hickner J. Barriers and motivators for making error reports from family medicine offices: a report from the American academy of family physicians national research network (AAFP NRN). J Am Board Fam Med 2007;20(2):115-123 [FREE Full text] [doi: 10.3122/jabfm.2007.02.060081] [Medline: 17341747]
23. Soydemir D, Seren Intepeler S, Mert H. Barriers to medical error reporting for physicians and nurses. West J Nurs Res 2017 Oct;39(10):1348-1363. [doi: 10.1177/0193945916671934] [Medline: 27694427]

XSL•FO
RenderX

24. Mirbaha F, Shalviri G, Yazdizadeh B, Gholami K, Majdzadeh R. Perceived barriers to reporting adverse drug events in hospitals: a qualitative study using theoretical domains framework approach. Implement Sci 2015 Aug 07;10:110 [FREE Full text] [doi: 10.1186/s13012-015-0302-5] [Medline: 26250159]
25. Mousavi-Roknabadi RS, Momennasab M, Groot G, Askarian M, Marjadi B. Medical error and under-reporting causes from the viewpoints of nursing managers: a qualitative study. Int J Prev Med 2022 Jul 15;13:103 [FREE Full text] [doi: 10.4103/ijpvm.IJPVM_500_20] [Medline: 36119955]
26. ISO 9241-11:2018 ergonomics of human-system interaction part 11: usability: definitions and concepts. International Organization for Standardization. URL: https://www.iso.org/standard/63500.html [accessed 2023-08-22]
27. Pot JW, de Vries CG. Medical devices in intramural settings and patient safety: availability of incident data and possible solutions to prevent use problems. National Institute for Public Health and the Environment. 2008. URL: https://www.rivm.nl/bibliotheek/rapporten/360020005.pdf [accessed 2024-01-11]
28. St-Maurice J, Burns C, Wolting J. Applying persuasive design techniques to influence data-entry behaviors in primary care: repeated measures evaluation using statistical process control. JMIR Hum Factors 2018 Oct 11;5(4):e28 [FREE Full text] [doi: 10.2196/humanfactors.9029] [Medline: 30309836]
29. Medication safety software for healthcare professionals. Pharmapod. URL: https://pharmapodhq.com/ [accessed 2023-08-21]
30. Oinas-Kukkonen H, Harjumaa M. Persuasive systems design: key issues, process model, and system features. Commun Assoc Inf Syst 2009;24 [FREE Full text] [doi: 10.17705/1CAIS.02428]
31. van Gemert-Pijnen L, Kelders SM, Beerlage-de Jong N, Oinas-Kukkonen H. Persuasive health technology. In: eHealth Research, Theory and Development. Milton Park, UK: Routledge; 2018.
32. Widyasari YD, Nugroho LE, Permanasari AE. Persuasive technology for enhanced learning behavior in higher education. Int J Educ Technol High Educ 2019 Apr 30;16:15. [doi: 10.1186/s41239-019-0142-5]
33. Stewart D, Thomas B, MacLure K, Wilbur K, Wilby K, Pallivalapila A, et al. Exploring facilitators and barriers to medication error reporting among healthcare professionals in Qatar using the theoretical domains framework: a mixed-methods approach. PLoS One 2018 Oct 2;13(10):e0204987 [FREE Full text] [doi: 10.1371/journal.pone.0204987] [Medline: 30278077]
34. Bullen K, Hall N, Sherwood J, Wake N, Donovan G. Prescribing error reporting in primary care: a narrative synthesis systematic review. Integr Healthc J 2020 Dec 10;2(1):e000026 [FREE Full text] [doi: 10.1136/ihj-2019-000026] [Medline: 37441317]
35. Holmström AR, Airaksinen M, Weiss M, Wuliji T, Chan XH, Laaksonen R. National and local medication error reporting systems: a survey of practices in 16 countries. J Patient Saf 2012 Dec;8(4):165-176. [doi: 10.1097/PTS.0b013e3182676cf3] [Medline: 23007239]
36. Alsaleh FM, Alsaeed S, Alsairafi ZK, Almandil NB, Naser AY, Bayoud T. Medication errors in secondary care hospitals in Kuwait: the perspectives of healthcare professionals. Front Med (Lausanne) 2021 Dec 20;8:784315 [FREE Full text] [doi: 10.3389/fmed.2021.784315] [Medline: 34988097]
37. Adams KT, Pruitt Z, Kazi S, Hettinger AZ, Howe JL, Fong A, et al. Identifying health information technology usability issues contributing to medication errors across medication process stages. J Patient Saf 2021 Dec 01;17(8):e988-e994. [doi: 10.1097/PTS.0000000000000868] [Medline: 34009868]
38. Cheung KC, van der Veen W, Bouvy ML, Wensing M, van den Bemt PM, de Smet PA. Classification of medication incidents associated with information technology. J Am Med Inform Assoc 2014 Feb;21(e1):e63-e70 [FREE Full text] [doi: 10.1136/amiajnl-2013-001818] [Medline: 24064444]
39. Marcilly R, Schiro J, Beuscart-Zéphir MC, Magrabi F. Building usability knowledge for health information technology: a usability-oriented analysis of incident reports. Appl Clin Inform 2019 May;10(3):395-408 [FREE Full text] [doi: 10.1055/s-0039-1691841] [Medline: 31189203]
40. Hautamäki E, Kinnunen UM, Palojoki S. Health information systems´ usability-related use errors in patient safety incidents. Finn J eHealth eWelfare 2017 Mar 31;9(1):6-17. [doi: 10.23996/fjhw.60763]
41. Kushniruk A, Patel V, Cimino JJ, Barrows RA. Cognitive evaluation of the user interface and vocabulary of an outpatient information system. Proc AMIA Annu Fall Symp 1996:22-26 [FREE Full text] [Medline: 8947620]
42. Ratwani RM, Savage E, Will A, Fong A, Karavite D, Muthu N, et al. Identifying electronic health record usability and safety challenges in pediatric settings. Health Aff (Millwood) 2018 Nov;37(11):1752-1759. [doi: 10.1377/hlthaff.2018.0699] [Medline: 30395517]
43. Kushniruk AW, Triola MM, Borycki EM, Stein B, Kannry JL. Technology induced error and usability: the relationship between usability problems and prescription errors when using a handheld application. Int J Med Inform 2005 Aug;74(7-8):519-526. [doi: 10.1016/j.ijmedinf.2005.01.003] [Medline: 16043081]
44. Fogg BJ. Persuasive Technology: Using Computers to Change What We Think and Do. Amsterdam, The Netherlands: Elsevier Science; 2003.
45. Oyibo K. Investigating the key persuasive features for fitness app design and extending the persuasive system design model: a qualitative approach. Proc Int Symp Hum Factors Ergon Health Care 2021 Jul 22;10(1):47-53 [FREE Full text] [doi: 10.1177/2327857921101022]

XSL•FO
RenderX

46. Bartlett YK, Webb TL, Hawley MS. Using persuasive technology to increase physical activity in people with chronic obstructive pulmonary disease by encouraging regular walking: a mixed-methods study exploring opinions and preference. J Med Internet Res 2017 Apr 20;19(4):e124 [FREE Full text] [doi: 10.2196/jmir.6616] [Medline: 28428155]
47. Win KT, Mullan J, Howard S, Oinas-Kukkonen H. Persuasive systems design features in promoting medication management for consumers. In: Proceedings of the 50th Hawaii International Conference on System Sciences. 2017 Presented at: 50th Hawaii International Conference on System Sciences; January 4-7, 2017; Hilton Waikoloa Village, Hawaii. [doi: 10.24251/hicss.2017.403]
48. Oyibo K, Vassileva J. HOMEX: persuasive technology acceptance model and the moderating effect of culture. Front Comput Sci 2020 Mar;2:1-17 [FREE Full text] [doi: 10.3389/fcomp.2020.00010]
49. TCPS 2 (2022) – Chapter 2: Scope and Approach. Government of Canada. URL: https://tinyurl.com/bdfhatrw [accessed 2023-01-25]
50. Do I need research ethics approval? University of Alberta. URL: https://tinyurl.com/mu8wkue4 [accessed 2024-01-25]
51. Lee BH, Lehmann CU, Jackson EV, Kost-Byerly S, Rothman S, Kozlowski L, et al. Assessing controlled substance prescribing errors in a pediatric teaching hospital: an analysis of the safety of analgesic prescription practice in the transition from the hospital to home. J Pain 2009 Feb;10(2):160-166 [FREE Full text] [doi: 10.1016/j.jpain.2008.08.004] [Medline: 19010736]
52. Calabrese AD, Erstad BL, Brandl K, Barletta JF, Kane SL, Sherman DS. Medication administration errors in adult patients in the ICU. Intensive Care Med 2001 Oct;27(10):1592-1598. [doi: 10.1007/s001340101065] [Medline: 11685299]
53. Cabri A, Barsegyan N, Postelnick M, Schulz L, Nguyen V, Szwak J, et al. Pharmacist intervention on prescribing errors: use of a standardized approach in the inpatient setting. Am J Health Syst Pharm 2021 Nov 23;78(23):2151-2158 [FREE Full text] [doi: 10.1093/ajhp/zxab278] [Medline: 34283219]
54. Boucher A, Ho C, MacKinnon N, Boyle TA, Bishop A, Gonzalez P, et al. Quality-related events reported by community pharmacies in Nova Scotia over a 7-year period: a descriptive analysis. CMAJ Open 2018 Dec 18;6(4):E651-E656 [FREE Full text] [doi: 10.9778/cmajo.20180090] [Medline: 30563921]
55. Stockton KR, Wickham ME, Lai S, Badke K, Villanyi D, Ho V, et al. LO51: incidence of clinically relevant medication errors after implementation of an electronic medication reconciliation process. Can J Emerg Med 2017 May 15;19(S1):S45. [doi: 10.1017/cem.2017.113]
56. Brennan-Bourdon LM, Vázquez-Alvarez AO, Gallegos-Llamas J, Koninckx-Cañada M, Marco-Garbayo JL, Huerta-Olvera SG. A study of medication errors during the prescription stage in the pediatric critical care services of a secondary-tertiary level public hospital. BMC Pediatr 2020 Dec 05;20(1):549 [FREE Full text] [doi: 10.1186/s12887-020-02442-w] [Medline: 33278900]
57. Shanks L, Bil K, Fernhout J. Learning without borders: a review of the implementation of medical error reporting in Médecins Sans Frontières. PLoS One 2015 Sep 18;10(9):e0137158 [FREE Full text] [doi: 10.1371/journal.pone.0137158] [Medline: 26381622]
58. Fyhr A, Akselsson R. Characteristics of medication errors with parenteral cytotoxic drugs. Eur J Cancer Care (Engl) 2012 Sep 16;21(5):606-613 [FREE Full text] [doi: 10.1111/j.1365-2354.2012.01331.x] [Medline: 22335508]
59. Jambrina AM, Santomà A, Rocher A, Rams N, Cereza G, Rius P, et al. Detection and prevention of medication errors by the network of sentinel pharmacies in a southern European region. J Clin Med 2022 Dec 27;12(1):194 [FREE Full text] [doi: 10.3390/jcm12010194] [Medline: 36614995]
60. Aljadhey H, Mahmoud MA, Ahmed Y, Sultana R, Zouein S, Alshanawani S, et al. Incidence of adverse drug events in public and private hospitals in Riyadh, Saudi Arabia: the (ADESA) prospective cohort study. BMJ Open 2016 Jul 12;6(7):e010831 [FREE Full text] [doi: 10.1136/bmjopen-2015-010831] [Medline: 27406640]
61. Al-Harkan A, Al-Harkan N, Al-Najjar A, Al-Hunti A, Al-Rashidi A, Al-Themery A. Investigation of medication errors in a tertiary care hospitals in the Qassim Region, Saudi Arabia. Open Access Maced J Med Sci 2020 Apr 20;8(E):209-212. [doi: 10.3889/oamjms.2020.4330]
62. Qu C, Meng L, Wang N, Chen Y, Yang X, Wang J, et al. Identify and categorize drug-related problems in hospitalized surgical patients in China. Int J Clin Pharm 2019 Feb 4;41(1):13-17. [doi: 10.1007/s11096-018-0777-7] [Medline: 30610549]
63. Han JH, Heo KN, Han J, Lee MS, Kim SJ, Min S, et al. Analysis of medication errors reported by community pharmacists in the Republic of Korea: a cross-sectional study. Medicina (Kaunas) 2023 Jan 12;59(1):151 [FREE Full text] [doi: 10.3390/medicina59010151] [Medline: 36676775]
64. Pharmapod homepage. Pharmapod. URL: https://ca.pharmapodhq.com/ [accessed 2023-08-21]
65. Karsh BT, Escoto KH, Beasley JW, Holden RJ. Toward a theoretical approach to medical error reporting system research and design. Appl Ergon 2006 May;37(3):283-295 [FREE Full text] [doi: 10.1016/j.apergo.2005.07.003] [Medline: 16182233]
66. Orji R, Oyibo K, Tondello GF. A comparison of system-controlled and user-controlled personalization approaches. In: Proceedings of the Adjunct Publication of the 25th Conference on User Modeling, Adaptation and Personalization. 2017 Presented at: UMAP '17; July 9-12, 2017; Bratislava, Slovakia. [doi: 10.1145/3099023.3099116]
67. Salam SN, Yahaya WA, Ali AM. Using persuasive design principles in motivational feeling towards children dental anxiety (CDA). In: Proceedings of the Persuasive Technology, 5th International Conference. 2010 Presented at: Persuasive 2010; June 7-10, 2010; Copenhagen, Denmark. [doi: 10.1007/978-3-642-13226-1_23]

XSL•FO
RenderX

68. Jenkins R, Vaida A. Simple strategies to avoid medication errors. Fam Pract Manag 2007 Feb;14(2):41-47 [FREE Full text] [Medline: 17330716]
69. Koczmara C, Dueck C, Jelincic V. To err is human, to share is divine. Dynamics 2006;17(3):22-25. [Medline: 17009570]
70. Reporting and learning systems for medication errors: the role of pharmacovigilance centres. World Health Organization. 2014 Feb 22. URL: https://www.who.int/publications-detail-redirect/9789241507943 [accessed 2023-08-22]
71. Nielsen J, Molich R. Heuristic evaluation of user interfaces. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. 1990 Presented at: CHI '90; April 1-5, 1990; Seattle, Washington. [doi: 10.1145/97243.97281]
72. Chen P. 7 elements of persuasive design strategy. Association for Talent Development. 2017 Dec 19. URL: https://www.td.org/insights/7-elements-of-persuasive-design-strategy [accessed 2023-08-21]
73. Cialdini RB. Influence, New and Expanded: The Psychology of Persuasion. New York, NY: HarperCollins Publishers; 2021.
74. Orji R, Lomotey R, Oyibo K, Orji F, Blustein J, Shahid S. Tracking feels oppressive and 'punishy': exploring the costs and benefits of self-monitoring for health and wellness. Digit Health 2018;4:2055207618797554 [FREE Full text] [doi: 10.1177/2055207618797554] [Medline: 30202544]
75. Cham S, Algashami A, McAlaney J, Stefanidis A, Phalp K, Ali R. Goal setting for persuasive information systems: five reference checklists. Lect Notes Comput Sci 2019;11433 LNCS:237-253. [doi: 10.1007/978-3-030-17287-9_20]
76. Selassie HH, Oyibo K, Vassileva J. Responsiveness to persuasive strategies at the workplace: a case study. In: Proceedings of the E-Technologies: Embracing the Internet of Things. 2017 Presented at: MCETECH 2017; May 17-19, 2017; Ottawa, ON. [doi: 10.1007/978-3-319-59041-7_17]
77. Oyibo K, Adaji I, Orji R, Olabenjo B, Vassileva J. Susceptibility to persuasive strategies: a comparative analysis of Nigerians vs. Canadians. In: Proceedings of the 26th Conference on User Modeling, Adaptation and Personalization. 2018 Presented at: UMAP '18; July 8-11, 2018; Singapore, Singapore URL: https://dl.acm.org/doi/10.1145/3209219.3209239 [doi: 10.1145/3209219.3209239]
78. Morton K, Sutton S, Hardeman W, Troughton J, Yates T, Griffin S, et al. A text-messaging and pedometer program to promote physical activity in people at high risk of type 2 diabetes: the development of the PROPELS follow-on support program. JMIR Mhealth Uhealth 2015 Dec 15;3(4):e105 [FREE Full text] [doi: 10.2196/mhealth.5026] [Medline: 26678750]
79. Oyibo K, Olagunju AH, Olabenjo B, Adaji I, Deters R, Vassileva J. BEN'FIT: design, implementation and evaluation of a culture-tailored fitness app. In: Proceedings of the Adjunct Publication of the 27th Conference on User Modeling, Adaptation and Personalization. 2019 Presented at: UMAP'19 Adjunct; June 9-12, 2019; Larnaca, Cyprus. [doi: 10.1145/3314183.3323854]
80. Matthews J, Win KT, Oinas-Kukkonen H, Freeman M. Persuasive technology in mobile applications promoting physical activity: a systematic review. J Med Syst 2016 Mar 9;40(3):72. [doi: 10.1007/s10916-015-0425-x] [Medline: 26748792]
81. Oyibo K. Designing culture-tailored persuasive technology to promote physical activity. University of Saskatchewan. 2020. URL: https://harvest.usask.ca/handle/10388/12943 [accessed 2023-08-22]
82. van Loon MH. Self-assessment and self-reflection to measure and improve self-regulated learning in the workplace. In: McGrath S, Mulder M, Papier J, Suart, R, editors. Handbook of Vocational Education and Training. Cham, Switzerland: Springer; 2018.
83. Berkovsky S, Freyne J, Oinas-Kukkonen H. Influencing individually: fusing personalization and persuasion. ACM Trans Interact Intell Syst 2012 Jun;2(2):1-8. [doi: 10.1145/2209310.2209312]
84. Mohadis HM, Ali NM. Designing persuasive application to encourage physical activity at workplace among older workers. In: Proceedings of the Sixth International Conference on Digital Information and Communication Technology and its Applications (DICTAP). 2016 Presented at: Sixth International Conference on Digital Information and Communication Technology and its Applications (DICTAP); July 21-23, 2016; Konya, Turkey. [doi: 10.1109/dictap.2016.7544013]
85. Orji R, Vassileva J, Mandryk RL. Modeling the efficacy of persuasive strategies for different gamer types in serious games for health. User Model User Adapt Inter 2014 Jul 14;24(5):453-498. [doi: 10.1007/s11257-014-9149-8]
86. Locke EA, Latham GP. Building a practically useful theory of goal setting and task motivation: a 35-year odyssey. Am Psychol 2002 Sep;57(9):705-717. [doi: 10.1037/0003-066x.57.9.705]
87. Schembre SM, Liao Y, Robertson MC, Dunton GF, Kerr J, Haffey ME, et al. Just-in-time feedback in diet and physical activity interventions: systematic review and practical design framework. J Med Internet Res 2018 Mar 22;20(3):e106 [FREE Full text] [doi: 10.2196/jmir.8701] [Medline: 29567638]
88. Kluger AN, DeNisi A. The effects of feedback interventions on performance: a historical review, a meta-analysis, and a preliminary feedback intervention theory. Psychol Bull 1996 Mar;119(2):254-284. [doi: 10.1037/0033-2909.119.2.254]
89. Special edition: tall man lettering; ISMP updates its list of drug names with tall man letters. Institute for Safe Medication Practices. 2016 Jun 2. URL: https://tinyurl.com/3skfpa2k [accessed 2023-08-22]
90. Langrial S. Exploring the influence of persuasive reminders and virtual rehearsal on the efficacy of health behavior change support systems. University of Oulu. 2014. URL: http://jultika.oulu.fi/files/isbn9789526204765.pdf [accessed 2024-01-11]
91. van Dantzig S, Geleijnse G, van Halteren AT. Toward a persuasive mobile application to reduce sedentary behavior. Pers Ubiquit Comput 2012 Jul 12;17(6):1237-1246. [doi: 10.1007/s00779-012-0588-0]

XSL•FO
RenderX

92. O'Leary K, Liu L, McClure JB, Ralston J, Pratt W. Persuasive reminders for health self-management. AMIA Annu Symp Proc 2016;2016:994-1003 [FREE Full text] [Medline: 28269896]
93. Hollas T, Lane M, Coyne J, Ellis C. What is verbal persuasion. In: Loureiro MJ, Loureiro A, Gerber HR, editors. Advances in Educational Technologies and Instructional Design. Hershey, PA: IGI Global; 2023:58-74.
94. Empowerment and change: the purpose of counselling. Pearson. URL: https://tinyurl.com/t5cxrw5r [accessed 2023-12-21]
95. Kim C, Jeon HG, Lee KC. Discovering the role of emotional and rational appeals and hidden heterogeneity of consumers in advertising copies for sustainable marketing. Sustainability 2020 Jun 25;12(12):5189. [doi: 10.3390/su12125189]
96. Oyibo K. The relationship between perceived health message motivation and social cognitive beliefs in persuasive health communication. Information 2021 Aug 28;12(9):350. [doi: 10.3390/info12090350]
97. Oyibo K, Ali YS, Vassileva J. An empirical analysis of the perception of mobile website interfaces and the influence of culture. In: Proceedings of the Personalization in Persuasive Technology Workshop, Persuasive Technology. 2016 Presented at: Personalization in Persuasive Technology Workshop, Persuasive Technology 2016; April 5, 2016; Salzburg, Austria.
98. Oyibo K, Vassileva J. Investigation of persuasive system design predictors of competitive behavior in fitness application: a mixed-method approach. Digit Health 2019 Nov 01;5:2055207619878601 [FREE Full text] [doi: 10.1177/2055207619878601] [Medline: 31700652]
99. Mcleod S. Albert Bandura's Social Learning Theory. SimplyPsychology. 2023 Oct 24. URL: https://www.simplypsychology.org/bandura.html [accessed 2023-08-22]
100. Lyons SD, Berge ZL. Social learning theory. In: Seel NM, editor. Encyclopedia of the Sciences of Learning. Boston, MA: Springer; 2012.
101. Festinger L. A theory of social comparison processes. Hum Relat 2016 Apr 22;7(2):117-140. [doi: 10.1177/001872675400700202]
102. Foster D, Linehan C, Kirman B, Lawson S, James G. Motivating physical activity at work: using persuasive social media for competitive step counting. In: Proceedings of the 14th International Academic MindTrek Conference: Envisioning Future Media Environments. 2010 Presented at: MindTrek '10; October 6-8, 2010; Tampere, Finland. [doi: 10.1145/1930488.1930510]
103. Oyibo K, Orji R, Vassileva J. The influence of culture in the effect of age and gender on social influence in persuasive technology. In: Proceedings of the Adjunct Publication of the 25th Conference on User Modeling, Adaptation and Personalization. 2017 Presented at: UMAP '17; July 9-12, 2017; Bratislava, Slovakia. [doi: 10.1145/3099023.3099071]
104. Myers DG. Social Psychology. New York, NY: McGraw-Hill Education; 2009.
105. Oyibo K, Morita PP. The effect of persuasive design on the adoption of exposure notification apps: quantitative study based on COVID alert. JMIR Form Res 2022 Sep 6;6(9):e34212 [FREE Full text] [doi: 10.2196/34212] [Medline: 35580138]
106. AIMS (Assurance and Improvement in Medication Safety). Ontario College of Pharmacists. URL: https://www.ocpinfo.com/regulations-standards/aims-assurance-and-improvement-in-medication-safety/ [accessed 2023-08-22]
107. Yuan CL, Moon H, Kim KH, Wang S, Yu X. Third-party organization endorsement impacts on perceived value and B2B customer loyalty. Ind Mark Manage 2020 Oct;90:221-230. [doi: 10.1016/j.indmarman.2020.07.021]
108. Fogg BJ, Tseng H. The elements of computer credibility. In: Proceedings of the SIGCHI conference on Human Factors in Computing Systems. 1999 Presented at: CHI '99; May 15-20, 1999; Pittsburgh, PA. [doi: 10.1145/302979.303001]
109. Fogg BJ, Marshall J, Laraki O, Osipovich A, Varma C, Fang N, et al. What makes web sites credible?: a report on a large quantitative study. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. 2001 Presented at: CHI '01; March 31-April 5, 2001; Seattle, WA. [doi: 10.1145/365024.365037]
110. Verifiability. Collins Dictionary. URL: https://www.collinsdictionary.com/dictionary/english/verifiability [accessed 2023-08-22]
111. Jones R. Information and persuasive speaking. In: Communication in the Real World: An Introduction to Communication Studies. Minneapolis, MN: University of Minnesota Libraries Publishing; 2016.
112. Wolf ZR, Hughes RG. Error reporting and disclosure. In: Patient Safety and Quality: An Evidence-Based Handbook for Nurses. Rockville, MD: Agency for Healthcare Research and Quality (US); 2008.
113. Orji R, Mandryk RL. Developing culturally relevant design guidelines for encouraging healthy eating behavior. Int J Hum Comput Stud 2014 Feb;72(2):207-223. [doi: 10.1016/j.ijhcs.2013.08.012]
114. Oyibo K, Orji R, Vassileva J. Developing culturally relevant design guidelines for encouraging physical activity: a social cognitive theory perspective. J Healthc Inform Res 2018 Dec 31;2(4):319-352 [FREE Full text] [doi: 10.1007/s41666-018-0026-9] [Medline: 35415413]
115. Kaldjian LC, Jones EW, Wu BJ, Forman-Hoffman VL, Levi BH, Rosenthal GE. Reporting medical errors to improve patient safety: a survey of physicians in teaching hospitals. Arch Intern Med 2008 Jan 14;168(1):40-46. [doi: 10.1001/archinternmed.2007.12] [Medline: 18195194]

Abbreviations

AIMS: Assurance and Improvement in Medication Safety
CQI: continuous quality improvement

XSL•FO
RenderX

IM: Incident Management
MI: medication incident
MIRLS: Medication Incident Reporting and Learning Systems
PSD: Persuasive System Design
RQ: research question
SMART: specific, measurable, achievable, relevant, and timebound
UX: user experience

Edited by A Kushniruk; submitted 30.07.22; peer-reviewed by C Hohl, R Asbjornsen; comments to author 29.04.23; revised version received 29.08.23; accepted 20.11.23; published 21.03.24.

Please cite as:
Oyibo K, Gonzalez PA, Ejaz S, Naheyan T, Beaton C, O'Donnell D, Barker JR
Exploring the Use of Persuasive System Design Principles to Enhance Medication Incident Reporting and Learning Systems: Scoping Reviews and Persuasive Design Assessment
JMIR Hum Factors 2024;11:e41557
URL: https://humanfactors.jmir.org/2024/1/e41557
doi:10.2196/41557
PMID:38512325

©Kiemute Oyibo, Paola A Gonzalez, Sarah Ejaz, Tasneem Naheyan, Carla Beaton, Denis O'Donnell, James R Barker. Originally published in JMIR Human Factors (https://humanfactors.jmir.org), 21.03.2024. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on https://humanfactors.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Clinical Decision Support Requirements for Ventricular Tachycardia Diagnosis Within the Frameworks of Knowledge and Practice: Survey Study

Zhao Hu[1*], MM; Min Wang[2*], MS; Si Zheng[2], MS; Xiaowei Xu[2], MS; Zhuxin Zhang[1], MS; Qiaoyue Ge[3], MS; Jiao Li[2], PhD; Yan Yao[1], PhD, MD

[1]Arrhythmia Center, Fuwai Hospital, Chinese Academy of Medical Sciences & Peking Union Medical College/National Center for Cardiovascular Diseases, Beijing, China

[2]Institute of Medical Information, Chinese Academy of Medical Sciences & Peking Union Medical College, Beijing, China

[3]West China School of Public Health, West China Fourth Hospital, Sichuan University, Chengdu, China

[*]these authors contributed equally

Corresponding Author:
Yan Yao, PhD, MD
Arrhythmia Center
Fuwai Hospital
Chinese Academy of Medical Sciences & Peking Union Medical College/National Center for Cardiovascular Diseases
Beilishi Road 167
Beijing, 100037
China
Phone: 86 10 88322401
Email: ianyao@263.net.cn

Abstract

Background: Ventricular tachycardia (VT) diagnosis is challenging due to the similarity between VT and some forms of supraventricular tachycardia, complexity of clinical manifestations, heterogeneity of underlying diseases, and potential for life-threatening hemodynamic instability. Clinical decision support systems (CDSSs) have emerged as promising tools to augment the diagnostic capabilities of cardiologists. However, a requirements analysis is acknowledged to be vital for the success of a CDSS, especially for complex clinical tasks such as VT diagnosis.

Objective: The aims of this study were to analyze the requirements for a VT diagnosis CDSS within the frameworks of knowledge and practice and to determine the clinical decision support (CDS) needs.

Methods: Our multidisciplinary team first conducted semistructured interviews with seven cardiologists related to the clinical challenges of VT and expected decision support. A questionnaire was designed by the multidisciplinary team based on the results of interviews. The questionnaire was divided into four sections: demographic information, knowledge assessment, practice assessment, and CDS needs. The practice section consisted of two simulated cases for a total score of 10 marks. Online questionnaires were disseminated to registered cardiologists across China from December 2022 to February 2023. The scores for the practice section were summarized as continuous variables, using the mean, median, and range. The knowledge and CDS needs sections were assessed using a 4-point Likert scale without a neutral option. Kruskal-Wallis tests were performed to investigate the relationship between scores and practice years or specialty.

Results: Of the 687 cardiologists who completed the questionnaire, 567 responses were eligible for further analysis. The results of the knowledge assessment showed that 383 cardiologists (68%) lacked knowledge in diagnostic evaluation. The overall average score of the practice assessment was 6.11 (SD 0.55); the etiological diagnosis section had the highest overall scores (mean 6.74, SD 1.75), whereas the diagnostic evaluation section had the lowest scores (mean 5.78, SD 1.19). A majority of cardiologists (344/567, 60.7%) reported the need for a CDSS. There was a significant difference in practice competency scores between general cardiologists and arrhythmia specialists (P=.02).

Conclusions: There was a notable deficiency in the knowledge and practice of VT among Chinese cardiologists. Specific knowledge and practice support requirements were identified, which provide a foundation for further development and optimization of a CDSS. Moreover, it is important to consider clinicians' specialization levels and years of practice for effective and personalized support.

XSL•FO
RenderX

(JMIR Hum Factors 2024;11:e55802) doi:10.2196/55802

KEYWORDS

clinical decision support system; requirements analysis; ventricular tachycardia; knowledge; clinical practice; questionnaires

Introduction

Sudden cardiac death (SCD) remains a significant public health issue, accounting for 50% of all cardiovascular deaths. The estimated annual incidences of SCD are 60 [1], 40.7 [2,3], and 36.8 [4] per 100,000 people in the United States, China, and Europe, respectively. Ventricular tachycardia (VT) is a major cause or precursor of SCD [5], which can be the initial or sole manifestation of diverse heart diseases [6,7]. VT diagnosis is challenging due to its similarity with some forms of supraventricular tachycardia, the complexity of clinical manifestations, heterogeneity of underlying diseases, and potential for life-threatening hemodynamic instability [6,8]. Diagnostic accuracy and timing are critical for patients with VT, as the stage of diagnosis determines the selection of treatment [9]. However, studies have revealed a substantial prevalence of misdiagnoses of VT [10-13], focusing on differential diagnosis between VT and supraventricular tachycardia. Although diagnostic error has been a challenge along the development of medicine, measuring diagnostic error can be difficult due to detection and reporting biases, with scarce reports indicating error rates of approximately 10%-15% [14]. We could not find additional estimates for the actual diagnostic error of VT; however, it is commonly acknowledged to represent a substantial challenge considering the complexity of the condition [9,15].

Diagnosis represents a complex cognitive process comprising a variety of different problem-solving tasks that are related to the clinical reasoning process, such as taking a medical history, forming a differential diagnosis, ordering examinations, and interpreting clinical findings [16]. The diagnostic process requires not only the retention of knowledge but also the judicious application of that knowledge at opportune moments, namely in clinical practice. A proper diagnosis of VT demands a great volume of knowledge. First, the clinician must be able to identify VT among the spectrum of wide QRS tachycardias by inspecting a list of electrocardiogram (ECG) features and comparing the findings to various diagnostic criteria or algorithms [17,18]. Once VT is identified by ECG interpretation, the next step is to diagnose the underlying diseases from a vast disease spectrum. This is a particularly challenging task, as any disease involving the myocardium can cause VT, such as coronary artery disease (CAD), all types of cardiomyopathies, myocarditis, inherited arrhythmia syndromes, autoimmune or inflammatory diseases, and others [7,9]. Moreover, translating the enormous body of knowledge into proper practice can be difficult [19], which is exacerbated by the fact that VT can cause stress to clinicians due to the probability of hemodynamic instability.

In response to this challenge, the clinical decision support system (CDSS) has emerged as a promising tool to augment the diagnostic capabilities of clinicians. Clinical decision support (CDS) is a process for enhancing health-related decisions with pertinent, organized clinical knowledge and patient information, thus advancing health care delivery [20]. Use of a CDSS can provide clinicians with situation-specific knowledge that aids in making critical clinical decisions such as risk assessment, diagnosis, prognosis, and selection of therapy [21]. A clinical diagnostic decision support system (DDSS) is a computer-based algorithm that assists a clinician with one or more component steps of the diagnostic process [22]. A DDSS is expected to receive relevant patient information and return outputs to assist with the problems the clinician has encountered in the diagnostic process, such as suggesting a likely diagnosis. Some well-known DDSSs such as ISABEL [23] and Dxplain [24] provide a diagnosis list, which can offer a solution to the challenges associated with VT diagnosis. Most CDSSs exhibit efficacy in a laboratory or experimental environment; however, relatively few such systems are being used at present and the rate of use in routine clinical practice is low [20,25-27]. Studies have identified the main barriers to the widespread adoption of CDSSs, including vague requirements, poor integration with the clinical workflow, low user acceptance or trust, and lack of transparency. Among these barriers, comprehensive user requirements engineering should be performed at the very beginning of development, which should be continued iteratively throughout the CDSS design-development-implementation life cycle [25,26,28,29]. To address this gap, several recent studies have aimed at elucidating the clinical requirements for an effective and usable CDSS in the context of specific fields or scenarios [30-34] with a variety of methods, including focus groups [30,35], a workshop [34], expert discussion with a literature review [36,37], semistructured interviews [31,34,35,38], writing user stories [39], and system evaluation [40]. Overall, most studies have adopted a user-centered approach with qualitative analysis.

To our best knowledge, although an artificial intelligence model was reported for predicting the in-hospital mortality of VT [8], no CDSS has been developed for VT diagnosis. A recent systemic review of cardiovascular CDSSs found that the complexity of the clinical management of cardiovascular disease itself was a barrier during implementation [27], which emphasizes the need for an authentic clinical requirements analysis. Accordingly, the objective of this study was to analyze the requirements for a VT diagnosis CDSS within the frameworks of knowledge, practice, and CDS needs.

Methods

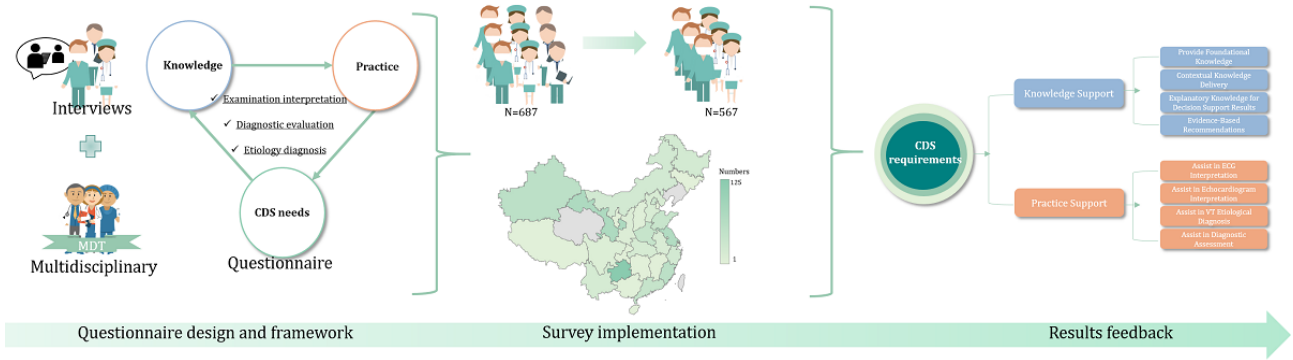
Study Design and Recruitment Process

Figure 1 shows the overall flow of our study, which consisted of semistructured interviews in the early stages and questionnaires in the later stages. To effectively implement and conduct the questionnaire assessment, we conducted open and explorative interviews about the challenges associated with the management of VT and the expected functions of a CDSS for VT. The interviews were conducted at Fuwai Hospital, the

XSL•FO
RenderX

national cardiovascular disease center of China. This hospital actively recruits cardiologists for their fellowships from all regions of China, resulting in a representative sample of interviewees. We sent interview invitations to all 56 cardiologists in the arrythmia center, including cardiologists from the fellowship program or established staff of Fuwai Hospital. Seven cardiologists responded and completed the interview, followed by a brief questionnaire to provide information on demographics and clinical experience (see Multimedia Appendix 1).

Figure 1. Schematic of the overall study workflow and assessment approach. CDS: clinical decision support; ECG: electrocardiogram; MDT: multidisciplinary team; VT: ventricular tachycardia.



A multidisciplinary team was formed to define the purpose of our study and the design of the questionnaire based on the interview results. The multidisciplinary team comprised three arrhythmia specialists, three experts in medical informatics and CDS, and one clinical statistician. The questionnaire was examined by an additional 20 arrhythmia specialists to ensure its clarity and feasibility. We conducted a nationwide cross-sectional survey with an online questionnaire in mainland China from December 31, 2022, to February 15, 2023. We recruited registered cardiologists using a convenience sampling approach from network groups associated with the Asian Heart Rhythm Association (AHRA) on WeChat, the dominant social media app in China. The AHRA is an academic organization focusing on arrhythmias, whose members are all registered cardiologists. Duplicate submissions were prevented through IP address constraints, and only completed responses were included for analysis.

Ethical Considerations

Participants provided online informed consent, which detailed the survey's background, aim, methods, and confidentiality measures. To protect participants' privacy, a signature was not required. Instead, participants clicked the "go on" button at the bottom of the informed consent page if they agreed to participate. According to data privacy protocols, no personal information, including the participants' names or affiliations, was collected. Since patients were not the subject of this study, ethical approval was exempted by the ethics committee of the Institute of Medical Information, Chinese Academy of Medical Sciences/Peking Union Medical College [41]. Each participant received ~US $3 as compensation.

Questionnaire Design

Overview

The questionnaire was divided into four sections (Table 1): demographic information (questions 1-6), knowledge assessment (question 14), practice assessment (questions 7-13), and CDS needs (questions 15-18). A comprehensive version of the questionnaire is provided in Multimedia Appendix 2.

Table 1. Design of the questionnaire.

Section	Content	Related questions
Knowledge	Examination interpretation, etiological diagnosis, diagnostic evaluation, conceptual knowledge	14
Practice	Examination interpretation, etiological diagnosis, diagnostic evaluation	7-13
Clinical decision support needs	Interpretable diagnosis, executable processes, knowledge support	15-18

Knowledge Assessment

Knowledge serves as the theoretical foundation for clinicians to make clinical diagnoses and is thus an essential competency for clinicians. The diagnosis of VT is difficult as it will largely depend on the clinician's familiarity with the vast knowledge of the field. The European Society of Cardiology (ESC) guideline suggests a protocol for VT diagnosis [15]. The multidisciplinary team abstracted the knowledge points from the ESC guideline for collecting information on the participants' self-reported knowledge shortcomings.

XSL•FO
RenderX

Practice Assessment

Areas of Focus

To attain a more accurate gauge of the clinical practice competency, we used simulated cases rather than straightforward questions [42], which can help differentiate practice competency from knowledge. To mitigate the risk of low response rates and careless submissions associated with lengthy surveys [43], we designed two stepwise cases containing seven questions. According to the intention, the questions about clinical practice were divided into three parts: examination interpretation, etiological diagnosis, and diagnostic evaluation. Multiple-choice options were available for all the questions. We standardized the total score for each section to 10 points according to the weighting.

Examination Interpretation

Accurate interpretation of an examination is the basis for a correct etiological diagnosis. ECG is the first-line examination modality for arrhythmias, as nearly all arrhythmia episodes are detected by ECG. Therefore, for this section, we focused on the identification of VT and sites of origin of VT on ECG [15].

Etiological Diagnosis

A correct etiological diagnosis of VT is necessary for appropriate treatment. The main strategy is to identify or exclude structural heart diseases, including CAD, myocarditis, and cardiomyopathies [44]. In this section, we assessed the correctness of a diagnosis of arrhythmogenic right ventricular cardiomyopathy (ARVC) and acute myocarditis as the two cases.

Diagnostic Evaluation

Diagnostic evaluation is a process of collecting clinical information to confirm or exclude a suspected diagnosis. A diagnostic evaluation protocol for VT is recommended in the ESC guideline [15] with the goal of reducing the rate of diagnostic errors. Based on the cases with an etiological diagnosis, we assessed the competency of the participants to arrange further diagnostic evaluations.

CDS Needs

According to the ESC guideline [15] and universal CDSS functionality [25], the multidisciplinary team summarized the results of the interviews to produce a list of functions required for CDS, which could be divided into executable processes, interpretable diagnosis, and knowledge support. We employed this list to poll the functionalities required by the cardiologists for a VT CDSS.

Quality Control of Responses

To ensure the validity and reliability of our survey responses, we used two strategies to filter out potentially low-quality submissions. First, participants who completed the questionnaire in under 2 minutes were excluded. This threshold was determined through a pretest evaluation coupled with multidisciplinary team discussions. Second, responses were considered to be invalid if participants selected all the available options for questions 7, 8, 9, 11, 12, or 13. This exclusion criterion was established based on the consensus opinion of the multidisciplinary team, who deemed such selections to be unreasonable.

Statistical Analysis

We only included valid questionnaire responses in the statistical analysis. All data in the demographic section were categorical. Comparisons were performed using mean, median, range, and percentage. The scores in the practice section are expressed as continuous variables, using the mean, median, and range. The knowledge and CDS sections were phrased as single-choice questions asking clinicians about their subjective views on given statements using a 4-point Likert scale without a neutral option. The internal consistency of the questionnaire was assessed using the Cronbach α value.

In addition, we grouped participants separately by practice years and specialty for further subgroup analyses. The Kruskal-Wallis test was performed to investigate the relationship between practice scores and practice years or specialty. All analyses were conducted in R version 4.0.3 [45]. We analyzed most of the data descriptively using graphics produced by the R package ggplot2.

Results

Sociodemographic Characteristics of Participants

A total of 687 questionnaires were completed. After applying our quality control measures, 567 responses were considered valid, yielding a validity rate of 82.53%. Among the invalid questionnaires, 104 responses were excluded due to a completion time of less than 2 minutes and 16 were excluded for selecting all options in questions 7, 8, 9, 11, 12, or 13. Descriptive statistics regarding the sociodemographic characteristics of participants are presented in Table 2. Of the enrolled participants, 54.50% were men; 93.47% were general cardiologists and the others were cardiac arrhythmia specialists. More than half of the participants were from tertiary A hospitals. Only a small percentage of cardiologists had ever used a CDSS, and the majority reported needing a CDSS to assist them in the management of VT (Table 2).

XSL•FO
RenderX

Table 2. Demographic characteristics of the survey participants (N=567).

Characteristics	Participants, n (%)
Gender	
Woman	258 (45.5)
Man	309 (54.50)
Age (years)	
≤30	89 (15.7)
31-35	152 (26.81)
36-40	129 (22.75)
41-45	92 (16.23)
46-50	60 (10.58)
≥51	45 (7.94)
Department	
Cardiology	530 (93.47)
Cardiac arrhythmia specialty	39 (6.88)
Professional title	
Resident physician	120 (21.16)
Attending	237 (41.8)
Associate chief	145 (25.57)
Chief	65 (11.46)
Years of practice	
<10	247 (43.54)
10-20	213 (37.57)
>20	107 (18.87)
Hospital tier	
Tertiary A	414 (73.02)
Not tertiary A	153 (26.98)
Ever used a CDSS[a]?	
Yes	72 (12.70)
No	495 (87.30)
Is there a need for a CDSS?	
Yes	523 (92.24)
No	44 (7.76)

[a]CDSS: clinical decision support system.

Semistructured Interviews

Textbox 1 summarizes the results of the semistructured interviews, in which we focused on the challenges of VT management and CDSS needs. The responses of the seven cardiologists were focused, with each noting that etiological diagnosis and interpretation of ECG results were their main challenges. The most important demand was the provision of quick and concise recommendations on diagnosis and treatment. The interviewees also expected the CDSS to provide clinical pathways.

XSL•FO
RenderX

Textbox 1. Results of the interviews.

- Challenges in the management of ventricular tachycardia (VT)

1. Etiological diagnosis
2. Wide QRS tachycardia diagnosis on electrocardiogram (ECG)
3. Determination of the location of VT origin on ECG
4. Mechanisms of VT
5. Drug treatment options
6. Options for the treatment of polymorphic VT

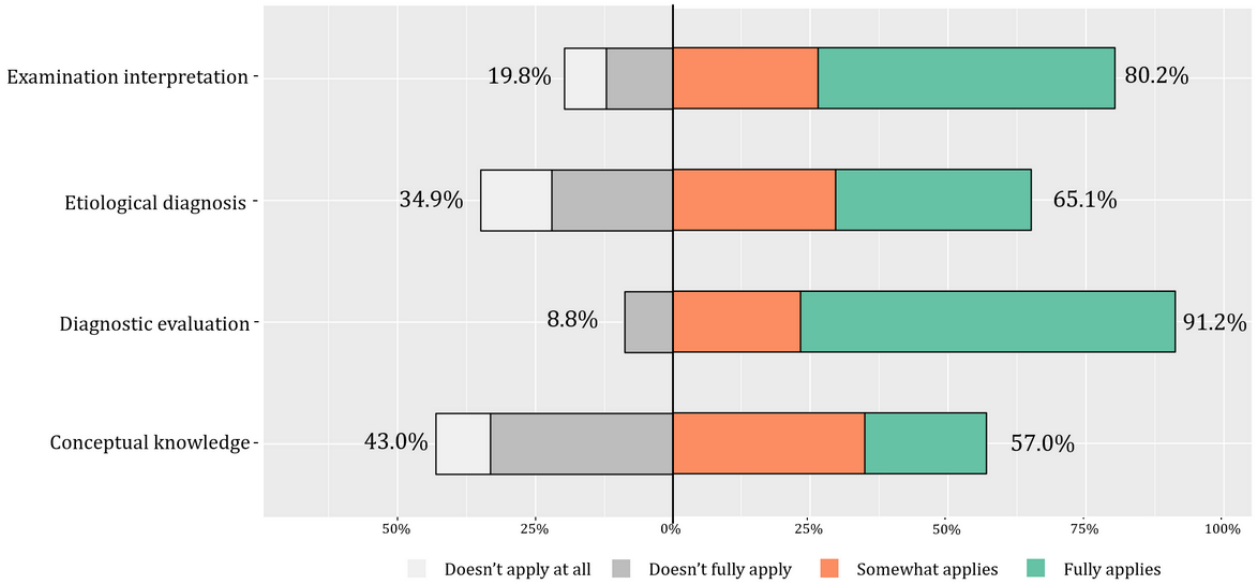
- Clinical decision support system needs

1. Rapid and concise recommendations for diagnosis and treatment
2. Diagnostic and therapeutic pathways for different etiologies
3. Aids in the identification of wide QRS
4. Adjunctive etiological diagnosis
5. Diagnostic supplements for related diseases

Knowledge

Figure 2 shows that there was an overall lack of knowledge with respect to diagnostic evaluation, with 383 of the 567 (68.0%) cardiologists indicating full need of assistant knowledge in diagnostic evaluation. This was followed by examination interpretation, where 305 of the 567 (53.8%) cardiologists were in full need of knowledge regarding the interpretation of ECG, cardiac ultrasound, and other cardiac examinations. The need for conceptual knowledge was relatively lower, even though it still reached nearly 60%.

Figure 2. Knowledge assessment.

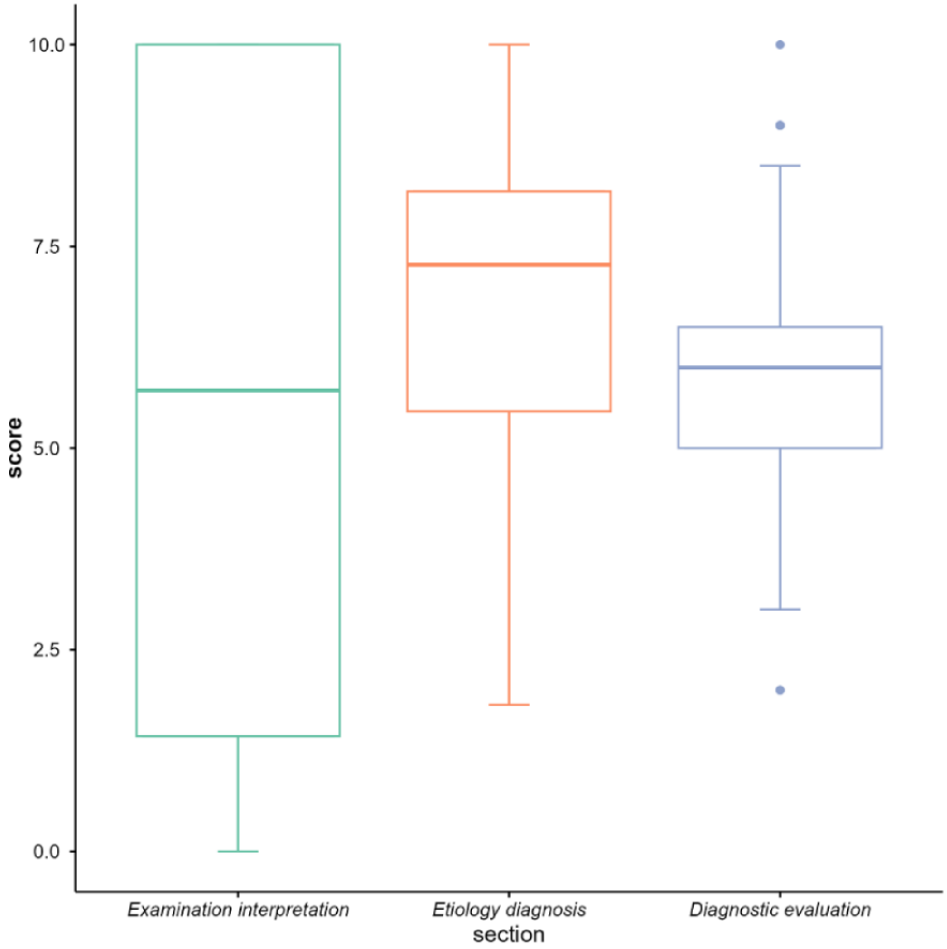


Practice

The overall average score of the practice questions was 6.11 (SD 0.55), the internal consistency of which was confirmed by a Cronbach α of 0.913. The mean scores of the examination interpretation, etiological diagnosis, and diagnostic evaluation were 6.22 (SD 3.94), 6.74 (SD 1.75), and 5.78 (SD 1.19), respectively. As shown in Figure 3, the etiological diagnosis section was associated with the highest overall score and the distribution of scores was also more concentrated than for the other sections, especially when compared with the distribution of the examination interpretation scores that were more dispersed and polarized.

XSL•FO
RenderX

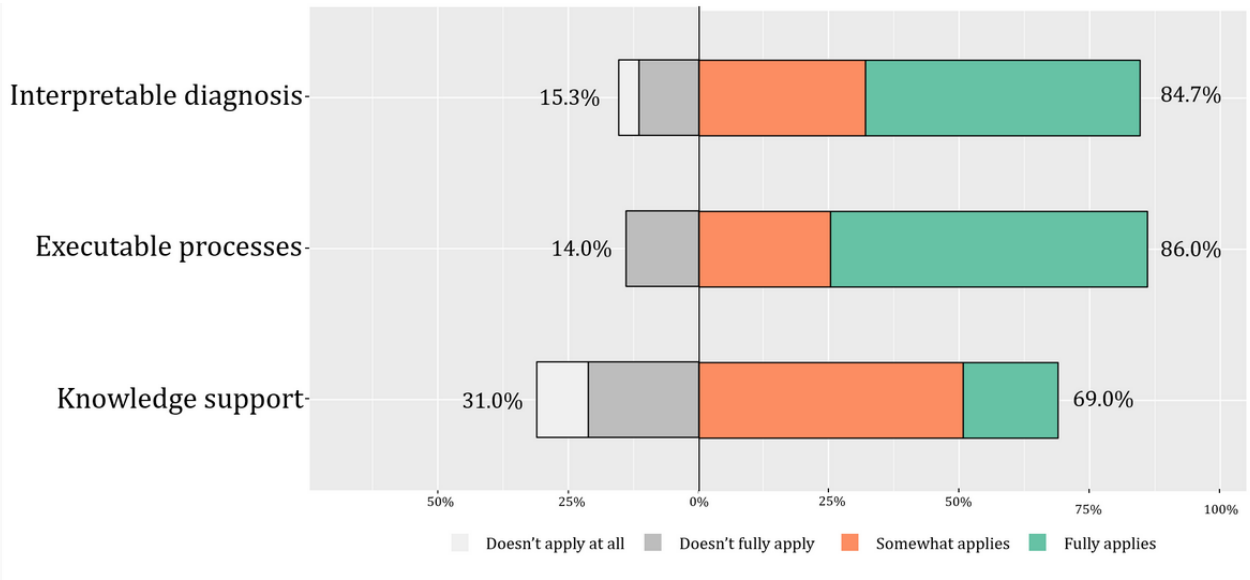
Figure 3. Practice assessment.



CDS Needs

The majority of the surveyed cardiologists reported a positive attitude toward CDS needs (Figure 4). There was relatively higher demand expressed for functions related to executable processes and interpretable diagnosis. In particular, the executable processes function was considered to be an essential requirement of a CDSS by 344 of the 567 cardiologists (60.7%). Knowledge support function received the least support but was still close to 70%.

Figure 4. Clinical decision support needs assessment.

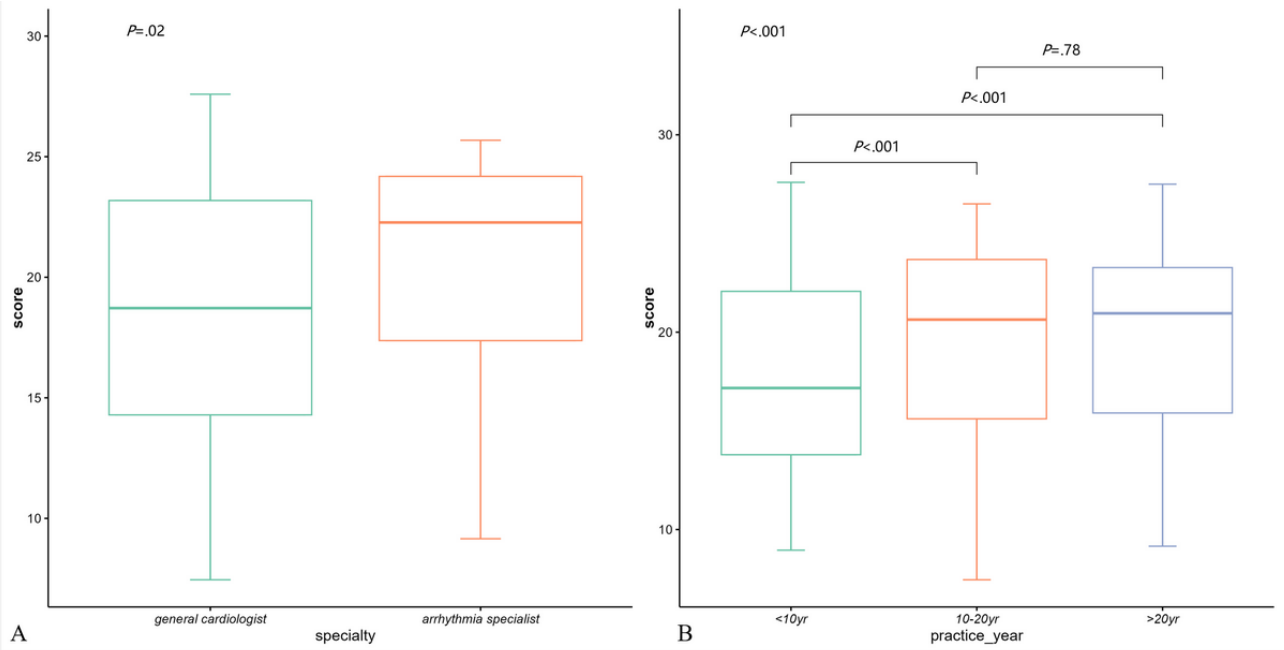


XSL•FO
RenderX

Subgroup Analysis

We divided all the cardiologists into subgroups based on specialty (Figure 5A) and practice years (Figure 5B). The Kruskal-Wallis test showed a significant difference in practice competency scores between general cardiologists and arrhythmia specialists (*P*=.02). Subgroup analysis according to years of practice revealed a significant effect of experience on scores. The <10 years group had significantly lower scores compared to those of the 10-20 years and >20 years groups. However, there was no significant difference between those with 10-20 years and >20 years of experience.

Figure 5. Subgroup analyses according to (A) specialty and (B) years of practice.



Discussion

Principal Results

Based on a combination of semistructured interviews and questionnaires, this study conducted a large-scale nationwide survey for cardiologists to understand their knowledge and practice competence about VT diagnosis and their requirements for a related CDSS. The results indicated that knowledge and practice support in examination interpretation, etiological diagnosis, and diagnostic evaluation are considered to be essential for a VT diagnosis CDSS. In addition, the vast majority of the cardiologists gave a positive response with respect to the need for a CDSS.

CDSS Requirements

Previous research on CDSS requirements has primarily relied on methods such as interviews [31,34,35,38,39] and group discussions [30,34,35] to elicit users' subjective needs. Based on recommendations from clinical experts and medical informatics professionals within our research team, it was acknowledged that certain objective requirements might not be articulated by users during interviews. Consequently, a questionnaire was designed to assess and uncover the requirements that might not have been spontaneously expressed during interviews. Previous studies have used questionnaires to investigate the knowledge, attitudes, and practices of health care professionals in various specific tasks [46-53], providing a basis for our questionnaire approach. To objectively reflect cardiologists' knowledge and practice deficiencies, we opted to not directly inquire about specific knowledge points but instead used two case scenarios to simulate authentic VT diagnostic situations, which is proven to be an appropriate method to assess practice competence [54]. The survey results endorsed the advantages of this mixed methods approach. The difficulties in VT diagnosis mentioned by the cardiologists during interviews primarily focused on distinguishing wide QRS tachycardias on ECG and identifying the etiology of VT, with no mention of diagnostic evaluation. However, results from the practice section of the questionnaire indicated poorer competence in diagnostic evaluation compared to etiological diagnosis, suggesting that the interviewees were not consciously aware of their weaknesses in diagnostic evaluation during interviews. Currently, there is no unified systematic method for conducting a CDS requirements analysis. While our method of integrating interviews and questionnaires provides a comprehensive approach, there is still room for improvement. Use of a simulation game has been suggested as a better means for clinical competence assessment [42]. Future research could consider incorporating cognitive analysis [55] and real-world system usability evaluation [56] to further optimize CDSS requirements analysis.

The objective results from case simulations also affirmed the cardiologists' need for decision support (Figure 4). Regarding knowledge requirements, the results from the CDS needs section of the questionnaire indicated that participants had relatively fewer demands for knowledge support compared to direct decision support. Moreover, the cardiologists revealed a preference for automatically prompted relevant knowledge during the diagnostic and therapeutic processes, which can provide more targeted knowledge support (Figure 2). The

XSL•FO
RenderX

challenge lies in ensuring that the CDSS accurately identifies the current diagnostic and therapeutic tasks; determines user knowledge gaps; and automatically retrieves, integrates, and presents knowledge support rapidly and accurately [57]. The results of the practice competence highlighted the need for improvement in the interpretation of diagnostic tests, etiological diagnosis, and diagnostic evaluation, suggesting the need for decision support in these three aspects, which were also highlighted as key clinical reasoning [58]. Notably, the accuracy of etiological diagnosis was relatively high, aligning with the lower knowledge demand for an etiological diagnosis (Figure 3). In terms of CDSS needs, the cardiologists favored direct decision support over knowledge support, including explanatory diagnoses and executable evaluation processes, which has also been recognized in recent studies [57,59,60].

Synthesizing the findings of this study, we propose the following recommendations of specific functions of a CDSS for VT diagnosis under a framework of knowledge and practice. With respect to knowledge support, the CDSS needs to (1) provide foundational knowledge by offering fundamental knowledge for each relevant disease that is available for clinicians to retrieve and browse; (2) contextualize knowledge delivery by providing closely related knowledge at decision points, including, but not limited to, the interpretation of diagnostic tests such as ECGs and echocardiograms, wide QRS complex differentiation, etiological diagnosis of VT, and the issuance of diagnostic test orders; (3) explain the knowledge underlying CDSS results; and (4) provide evidence-based recommendations at decision points with available evidence support. With respect to practice support, the CDDS should (1) assist in ECG interpretation, including distinguishing wide QRS complex tachycardias, identifying useful features for etiological diagnosis during sinus rhythm and VT, and recommending diagnostic test orders; (2) assist in echocardiogram interpretation, including the recognition of common etiologies of VT such as old myocardial infarction, ARVC, myocarditis, and the classification of phenotypes of cardiomyopathies; (3) provide suspected etiological diagnoses based on existing information for patients with VT, including acute coronary syndrome, ischemic cardiomyopathy, ARVC, and acute myocarditis, with specific emphasis on alerting clinicians who may not have considered the possibility of acute coronary syndrome; and (4) supplement diagnostic assessments with additional information, including critical medical history, physical examination, laboratory tests, and other examinations. Particularly, using a comprehensive differential diagnosis list is advocated to mitigate premature closure [14], as substantiated by a recent study [61].

Dxplain [24], one of the few DDSSs available for general practice, provides a diagnosis list according to input patient manifestations, which aligns with our proposed structure for VT etiological diagnosis. However, Dxplain lacks knowledge support, examination interpretation, and diagnostic assessment functions, which are highlighted as requirements for a VT CDSS as mentioned above. Another well-known commercial diagnostic support tool, ISABEL, not only serves as a diagnosis reminder but also provides knowledge support (ie, evidence-based knowledge of each disease). However, it does not satisfy the other requirements identified in this study [23,62]. Dr. Mayson [63] is a Chinese commercial CDSS for general practice, which can abstract data from electronic health records to form a diagnosis list as well as provide assistance in diagnostic assessment. Like ISABEL, Dr. Mayson provides a knowledge database for each disease, including clinical practice guidelines. However, the knowledge support is at the disease level rather than the decision level. In addition, this CDSS does not assist with examination interpretation.

Although our study mainly investigated the specific functionalities for VT diagnosis, the results indicated some general CDSS functionalities, including interpretability of decision-making as well as the overall feasibility of the CDSS workflow. Several reviews [64-66] summarized other universal features worthy of consideration, such as integration with the clinical workflow and electronic health record system, reduction of manual input of patient data, execution users' desired action, avoidance of unnecessary alerts, documentation of reasons for rejecting recommendations, as well as the "five rights" of CDS (providing the right information to the right people in the right formats through the right channels at the right time) [67].

We believe that an excellent CDSS should provide tailored assistance for different types of clinicians. Thus, a subgroup analysis was performed according to the clinician characteristics in the practice section (Figure 5). As anticipated, arrhythmia specialists outperformed general cardiologists, which aligns with the findings of previous research [68]. The American College of Cardiology defines different types of cardiovascular specialists that have requirements for different types of support in cardiovascular health care [69]. A CDSS should be tailored to clinicians' specialization levels to assist in diagnostic and therapeutic practices. For highly specialized clinicians facing a narrow spectrum of diseases, CDSS assistance may be limited, while support for foundational diagnostic and therapeutic aspects outside their specialty may be necessary. Conversely, less specialized clinicians facing a broader spectrum of diseases may need support in staying updated with the latest diagnostic and therapeutic advancements. For instance, for less experienced clinicians facing patients with VT, the CDSS should always indicate the possibility of CAD. For experienced clinicians, as they have already cultivated the mindset to exclude CAD, the CDSS might only provide this alert when they miss the diagnosis of CAD. Furthermore, it is expected that the CDSS could continually adapt to individual needs through observing clinician users' behaviors. The impact of years of practice on performance seems to be nonlinear. Clinicians practicing for 10-20 years or more demonstrated better performance than those practicing for less than 10 years. However, there was no significant difference between the 10-20 years and >20 years groups, suggesting that clinical skills may grow in the first 10 years of practice but plateau afterward, thereby challenging the CDSS design to provide targeted support for clinicians with different levels of experience in practice. Additionally, for clinicians entering a bottleneck period in competence growth, the CDSS could facilitate education during practice, thereby supporting lifelong learning. Several studies have been performed in this regard in the areas of pharmaceutical skills [70], imaging interpretation [71], geriatric care [72], and periprocedural antithrombotic use [73].

XSL•FO
RenderX

Most existing CDSSs have been generally designed for health care providers but might not fully consider the diversity of requirements as well as their expertise levels [74]. The genuine needs of health care providers have not been effectively communicated to system developers, resulting in the design of CDSSs that struggle to fulfill their intended role of assistance and workload reduction. Our study centers around the clinical scenario of VT diagnosis, comprehensively exploring support requirements in both knowledge and practice. This investigation can thus provide a foundation for the development of a relevant CDSS. Additionally, we aspire for this study to serve as a reference for clinical needs research, encouraging more health care providers and system developers to scrutinize clinical requirements and establish a groundwork for the development of highly effective CDSSs.

Limitations

Although this study used a combination of structured interviews and questionnaires for assessment, inevitably, some subjective factors from the participants may have biased the results. The questionnaire content of this study was carefully designed based on the results of the interviews as well as the experience of the multidisciplinary team; however, the questionnaire content was unable to cover all aspects of knowledge and practice related to VT diagnosis. Although specific functions for a VT diagnosis CDSS were proposed, they have not been evaluated in a real-world setting. As our team is currently developing a VT CDSS with these functions, more rigorous studies will be conducted to support these findings in our future research.

Conclusions

This comprehensive analysis of VT CDSS requirements using a mixed methods approach identified specific knowledge and practice support requirements. The derived functions provide a foundation for further development and optimization of a CDSS. Moreover, it is important to tailor the CDSS to clinicians' specialization levels and years of practice for effective and personalized support.

Acknowledgments

This study was supported by National High Level Hospital Clinical Research Funding (grant 2022-GSP-GG-25), the Chinese Academy of Medical Sciences Innovation Fund for Medical Sciences (CIFMS) (grant 2021-I2M-1-056), the Key R&D Program of Hunan Province (grant 2021SK2024), the Key Laboratory of Medical Information Intelligent Technology, and the National Key R&D Program of China (grants 2016YFC0901901 and 2017YFC0907503).

Data Availability

The data sets generated and/or analyzed during this study are available from the corresponding author on reasonable request.

Authors' Contributions

ZH, MW, SZ, XX, JL, and YY designed the study. ZH, MW, ZZ, YY, SZ, XX, JL, and QG designed the questionnaire. ZH and ZZ collected the data. MW, ZH, and QG analyzed the data. ZH and MW drafted the manuscript. SZ, XX, JL, and YY critically revised the manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Demographic characteristics and clinical experience of the interviewees.
[XLSX File (Microsoft Excel File), 9 KB - humanfactors_v11i1e55802_app1.xlsx]

Multimedia Appendix 2
Complete version of the questionnaire with the participant consent form.
[DOCX File , 979 KB - humanfactors_v11i1e55802_app2.docx]

References

1. Stecker EC, Reinier K, Marijon E, Narayanan K, Teodorescu C, Uy-Evanado A, et al. Public health burden of sudden cardiac death in the United States. Circ Arrhythm Electrophysiol 2014 Apr;7(2):212-217 [FREE Full text] [doi: 10.1161/CIRCEP.113.001034] [Medline: 24610738]
2. Feng X, Hai J, Ma Y, Wang Z, Tse H. Sudden cardiac death in mainland China: a systematic analysis. Circ Arrhythm Electrophysiol 2018 Nov;11(11):e006684. [doi: 10.1161/CIRCEP.118.006684] [Medline: 30571181]
3. Hua W, Zhang L, Wu Y, Liu X, Guo D, Zhou H, et al. Incidence of sudden cardiac death in China: analysis of 4 regional populations. J Am Coll Cardiol 2009 Sep 15;54(12):1110-1118 [FREE Full text] [doi: 10.1016/j.jacc.2009.06.016] [Medline: 19744622]

XSL•FO
RenderX

4. Empana J, Lerner I, Valentin E, Folke F, Böttiger B, Gislason G, ESCAPE-NET Investigators. Incidence of sudden cardiac death in the European Union. J Am Coll Cardiol 2022 May 10;79(18):1818-1827 [FREE Full text] [doi: 10.1016/j.jacc.2022.02.041] [Medline: 35512862]
5. Guandalini GS, Liang JJ, Marchlinski FE. Ventricular tachycardia ablation: past, present, and future perspectives. JACC Clin Electrophysiol 2019 Dec;5(12):1363-1383 [FREE Full text] [doi: 10.1016/j.jacep.2019.09.015] [Medline: 31857035]
6. Libby P, Bonow RO, Mann DL, Tomaselli GF, Bhatt DL, Solomon SD, et al. Braunwald's Heart Disease: A Textbook of Cardiovascular Medicine, 12th edition. Amsterdam: Elsevier; 2021.
7. Roberts-Thomson KC, Lau DH, Sanders P. The diagnosis and management of ventricular arrhythmias. Nat Rev Cardiol 2011 Jun 22;8(6):311-321. [doi: 10.1038/nrcardio.2011.15] [Medline: 21343901]
8. Li L, Ding L, Zhang Z, Zhou L, Zhang Z, Xiong Y, et al. Development and validation of machine learning-based models to predict in-hospital mortality in life-threatening ventricular arrhythmias: retrospective cohort study. J Med Internet Res 2023 Nov 15;25:e47664 [FREE Full text] [doi: 10.2196/47664] [Medline: 37966870]
9. Hayashi M, Shimizu W, Albert CM. The spectrum of epidemiology underlying sudden cardiac death. Circ Res 2015 Jun 05;116(12):1887-1906 [FREE Full text] [doi: 10.1161/CIRCRESAHA.116.304521] [Medline: 26044246]
10. Dancy M, Camm A, Ward D. Misdiagnosis of chronic recurrent ventricular tachycardia. Lancet 1985 Aug 10;2(8450):320-323. [doi: 10.1016/s0140-6736(85)90363-0] [Medline: 2862480]
11. Spice C, Stewart K. Misdiagnosis of ventricular tachycardia. Lancet 1999;354(9196):2165. [doi: 10.1016/S0140-6736(05)77068-9] [Medline: 10609843]
12. Stewart K. Misdiagnosis of ventricular tachycardia. Lancet 1986 Nov 08;2(8515):1103. [doi: 10.1016/s0140-6736(86)90506-4] [Medline: 2877260]
13. Morady F, Baerman JM, DiCarlo LA, DeBuitleir M, Krol RB, Wahr DW. A prevalent misconception regarding wide-complex tachycardias. JAMA 1985 Nov 15;254(19):2790-2792. [Medline: 4057488]
14. Graber ML, Franklin N, Gordon R. Diagnostic error in internal medicine. Arch Intern Med 2005 Jul 11;165(13):1493-1499. [doi: 10.1001/archinte.165.13.1493] [Medline: 16009864]
15. Zeppenfeld K, Tfelt-Hansen J, de Riva M, Winkel BG, Behr ER, Blom NA, ESC Scientific Document Group. 2022 ESC Guidelines for the management of patients with ventricular arrhythmias and the prevention of sudden cardiac death. Eur Heart J 2022 Oct 21;43(40):3997-4126. [doi: 10.1093/eurheartj/ehac262] [Medline: 36017572]
16. Denekamp Y, Peleg M. TiMeDDx--a multi-phase anchor-based diagnostic decision-support model. J Biomed Inform 2010 Feb;43(1):111-124 [FREE Full text] [doi: 10.1016/j.jbi.2009.08.001] [Medline: 19665579]
17. Abedin Z. Differential diagnosis of wide QRS tachycardia: a review. J Arrhythm 2021 Oct 09;37(5):1162-1172 [FREE Full text] [doi: 10.1002/joa3.12599] [Medline: 34621415]
18. Katritsis DG, Brugada J. Differential diagnosis of wide QRS tachycardias. Arrhythm Electrophysiol Rev 2020 Nov 05;9(3):155-160 [FREE Full text] [doi: 10.15420/aer.2020.20] [Medline: 33240511]
19. Ray S, Jeremy R, Nishimura R, Simpson IA. Identifying excellence in contemporary cardiology practice: transparency, professionalism and the role of the professional society. Heart 2013 Aug 23;99(16):1144-1145. [doi: 10.1136/heartjnl-2012-303483] [Medline: 23343687]
20. Osheroff J, Teich J, Levick D, Saldana L, Velasco F, Sittig D, et al. Improving Outcomes With Clinical Decision Support: An Implementer's Guide. 2nd ed. Milton Park: Routledge; 2012.
21. Osheroff JA, Teich JM, Middleton B, Steen EB, Wright A, Detmer DE. A roadmap for national action on clinical decision support. J Am Med Inform Assoc 2007;14(2):141-145 [FREE Full text] [doi: 10.1197/jamia.M2334] [Medline: 17213487]
22. Miller RA. Diagnostic decision support systems. In: Berner ES, editor. Clinical Decision Support Systems: Theory and Practice. Cham: Springer International Publishing; 2016:181-208.
23. ISABEL. URL: https://www.isabelhealthcare.com/ [accessed 2024-01-30]
24. DxPlain. mghls. URL: http://www.mghlcs.org/projects/dxplain [accessed 2024-01-30]
25. Greenes RA, Bates DW, Kawamoto K, Middleton B, Osheroff J, Shahar Y. Clinical decision support models and frameworks: seeking to address research issues underlying implementation successes and failures. J Biomed Inform 2018 Feb;78:134-143 [FREE Full text] [doi: 10.1016/j.jbi.2017.12.005] [Medline: 29246790]
26. Khairat S, Marc D, Crosby W, Al Sanousi A. Reasons for physicians not adopting clinical decision support systems: critical analysis. JMIR Med Inform 2018 Apr 18;6(2):e24 [FREE Full text] [doi: 10.2196/medinform.8912] [Medline: 29669706]
27. Lu Y, Melnick ER, Krumholz HM. Clinical decision support in cardiovascular medicine. BMJ 2022 May 25;377:e059818. [doi: 10.1136/bmj-2020-059818] [Medline: 35613721]
28. Timpka T, Johansson M. The need for requirements engineering in the development of clinical decision-support systems: a qualitative study. Methods Inf Med 1994 May;33(2):227-233. [Medline: 8057951]
29. Alvarez R. Discourse analysis of requirements and knowledge elicitation interviews. 2002 Presented at: 35th Annual Hawaii International Conference on System Sciences; January 7-10, 2002; Big Island, HI. [doi: 10.1109/hicss.2002.994388]
30. Bright T. Transforming user needs into functional requirements for an antibiotic clinical decision support system: explicating content analysis for system design. Appl Clin Inform 2013;4(4):618-635 [FREE Full text] [doi: 10.4338/ACI-2013-08-RA-0058] [Medline: 24454586]

XSL•FO
RenderX

31. Schüttler C, Hinderer M, Kraus S, Lang A, Prokosch H, Castellanos I. Requirements analysis for a clinical decision support system aiming at improving the artificial nutrition of critically ill patients. Stud Health Technol Inform 2017;243:137-141. [Medline: 28883187]
32. Miller K, Mosby D, Capan M, Kowalski R, Ratwani R, Noaiseh Y, et al. Interface, information, interaction: a narrative review of design and functional requirements for clinical decision support. J Am Med Inform Assoc 2018 May 01;25(5):585-592 [FREE Full text] [doi: 10.1093/jamia/ocx118] [Medline: 29126196]
33. Jones EK, Banks A, Melton GB, Porta CM, Tignanelli CJ. Barriers to and facilitators for acceptance of comprehensive clinical decision support system-driven care maps for patients with thoracic trauma: interview study among health care providers and nurses. JMIR Hum Factors 2022 Mar 16;9(1):e29019 [FREE Full text] [doi: 10.2196/29019] [Medline: 35293873]
34. Schütze D, Holtz S, Neff MC, Köhler SM, Schaaf J, Frischen LS, et al. Requirements analysis for an AI-based clinical decision support system for general practitioners: a user-centered design process. BMC Med Inform Decis Mak 2023 Jul 31;23(1):144 [FREE Full text] [doi: 10.1186/s12911-023-02245-w] [Medline: 37525175]
35. Kabukye J, Koch S, Cornet R, Orem J, Hagglund M. User requirements for an electronic medical records system for oncology in developing countries: a case study of Uganda. AMIA Annu Symp Proc 2017;2017:1004-1013 [FREE Full text] [Medline: 29854168]
36. Chaffee B. Developing and assessing requirements for clinical decision support. Am J Health Syst Pharm 2003 Sep 15;60(18):1875-1879. [doi: 10.1093/ajhp/60.18.1875] [Medline: 14521039]
37. Shah H, Allard RD, Enberg R, Krishnan G, Williams P, Nadkarni PM. Requirements for guidelines systems: implementation challenges and lessons from existing software-engineering efforts. BMC Med Inform Decis Mak 2012 Mar 09;12(1):16 [FREE Full text] [doi: 10.1186/1472-6947-12-16] [Medline: 22405400]
38. Gulla J, Neri PM, Bates DW, Samal L. User requirements for a chronic kidney disease clinical decision support tool to promote timely referral. Int J Med Inform 2017 May;101:50-57 [FREE Full text] [doi: 10.1016/j.ijmedinf.2017.01.018] [Medline: 28347447]
39. Kannan V, Basit MA, Bajaj P, Carrington AR, Donahue IB, Flahaven EL, et al. User stories as lightweight requirements for agile clinical decision support development. J Am Med Inform Assoc 2019 Nov 01;26(11):1344-1354 [FREE Full text] [doi: 10.1093/jamia/ocz123] [Medline: 31512730]
40. Balatsoukas P, Williams R, Davies C, Ainsworth J, Buchan I. User interface requirements for web-based integrated care pathways: evidence from the evaluation of an online care pathway investigation tool. J Med Syst 2015 Nov 7;39(11):183. [doi: 10.1007/s10916-015-0357-5] [Medline: 26446014]
41. Notice on issuing ethical review measures for life sciences and medical research involving humans. National Health Commission, Ministry of Education, Ministry of Science and Technology, Bureau of Traditional Chinese Medicine. URL: https://www.gov.cn/zhengce/zhengceku/2023-02/28/content_5743658.htm [accessed 2024-03-20]
42. Fonteneau T, Billion E, Abdoul C, Le S, Hadchouel A, Drummond D. Simulation game versus multiple choice questionnaire to assess the clinical competence of medical students: prospective sequential trial. J Med Internet Res 2020 Dec 16;22(12):e23254 [FREE Full text] [doi: 10.2196/23254] [Medline: 33325833]
43. Bradburn NM, Sudman S, Wansink B. Asking Questions: The Definitive Guide to Questionnaire Design--For Market Research, Political Polls, and Social and Health Questionnaires. Hoboken, NJ: John Wiley & Sons; 2004.
44. Shivkumar K, Do DH. Catheter ablation of ventricular tachycardia: first, treat the underlying disease. J Am Coll Cardiol 2020 Oct 06;76(14):1657-1659 [FREE Full text] [doi: 10.1016/j.jacc.2020.08.046] [Medline: 33004130]
45. R: A language and environment for statistical computing (Version 4.0.3). URL: https://www.R-project.org/ [accessed 2023-10-21]
46. Zhang C, Shen L, Pan M, Zheng Y, Gu Z, Lin H. Perceptions and knowledge gaps on CHADS-VASc score components: a joint survey of Chinese clinicians and clinical pharmacists. Postgrad Med 2022 Jan 08;134(1):64-77. [doi: 10.1080/00325481.2021.1996815] [Medline: 34694951]
47. Azraii AB, Ramli AS, Ismail Z, Abdul-Razak S, Badlishah-Sham SF, Mohd-Kasim NA, et al. Validity and reliability of an adapted questionnaire measuring knowledge, awareness and practice regarding familial hypercholesterolaemia among primary care physicians in Malaysia. BMC Cardiovasc Disord 2021 Jan 19;21(1):39 [FREE Full text] [doi: 10.1186/s12872-020-01845-y] [Medline: 33468051]
48. Malhotra S, Masood I, Giglio N, Pruetz JD, Pannaraj PS. Current knowledge of Chagas-related heart disease among pediatric cardiologists in the United States. BMC Cardiovasc Disord 2021 Mar 02;21(1):116. [doi: 10.1186/s12872-021-01924-8]
49. Montagni I, Roussel N, Thiébaut R, Tzourio C. Health care students' knowledge of and attitudes, beliefs, and practices toward the French COVID-19 app: cross-sectional questionnaire study. J Med Internet Res 2021 Mar 03;23(3):e26399 [FREE Full text] [doi: 10.2196/26399] [Medline: 33566793]
50. Sturny L, Regard S, Larribau R, Niquille M, Savoldelli GL, Sarasin F, et al. Differences in basic life support knowledge between junior medical students and lay people: web-based questionnaire study. J Med Internet Res 2021 Feb 23;23(2):e25125 [FREE Full text] [doi: 10.2196/25125] [Medline: 33620322]

XSL•FO
RenderX

51. Sadarmin PP, Wong KCK, Rajappan K, Bashir Y, Betts TR. Questionnaire survey of cardiologists' knowledge, attitudes, and guideline application of implantable cardioverter defibrillator therapy. Pacing Clin Electrophysiol 2012 Jun 27;35(6):672-680. [doi: 10.1111/j.1540-8159.2012.03373.x] [Medline: 22452348]
52. Bergman HE, Reeve BB, Moser RP, Scholl S, Klein WM. Development of a comprehensive heart disease knowledge questionnaire. Am J Health Educ 2011 Mar 23;42(2):74-87 [FREE Full text] [doi: 10.1080/19325037.2011.10599175] [Medline: 21720571]
53. Shah S, Davies M, Quinn T. Management of acute coronary syndromes, a questionnaire survey of the clinical practice of cardiologists and other medical physicians belonging to west midland hospitals. Int J Cardiol 2005 Mar 10;99(1):71-75. [doi: 10.1016/j.ijcard.2003.11.043] [Medline: 15721502]
54. Wass V, Van der Vleuten C, Shatzer J, Jones R. Assessment of clinical competence. Lancet 2001 Mar 24;357(9260):945-949. [doi: 10.1016/S0140-6736(00)04221-5] [Medline: 11289364]
55. Patel VL, Arocha JF, Diermeier M, Greenes RA, Shortliffe EH. Methods of cognitive analysis to support the design and evaluation of biomedical systems: the case of clinical practice guidelines. J Biomed Inform 2001 Feb;34(1):52-66 [FREE Full text] [doi: 10.1006/jbin.2001.1002] [Medline: 11376543]
56. Trinkley KE, Kahn MG, Bennett TD, Glasgow RE, Haugen H, Kao DP, et al. Integrating the practical robust implementation and sustainability model with best practices in clinical decision support design: implementation science approach. J Med Internet Res 2020 Oct 29;22(10):e19676 [FREE Full text] [doi: 10.2196/19676] [Medline: 33118943]
57. Mittermaier M, Raza M, Kvedar JC. Collaborative strategies for deploying AI-based physician decision support systems: challenges and deployment approaches. NPJ Digit Med 2023 Aug 05;6(1):137. [doi: 10.1038/s41746-023-00889-6] [Medline: 37543707]
58. Brush JE, Sherbino J, Norman GR. Diagnostic reasoning in cardiovascular medicine. BMJ 2022 Jan 05;376:e064389. [doi: 10.1136/bmj-2021-064389] [Medline: 34987062]
59. Garg AX, Adhikari NKJ, McDonald H, Rosas-Arellano MP, Devereaux PJ, Beyene J, et al. Effects of computerized clinical decision support systems on practitioner performance and patient outcomes: a systematic review. JAMA 2005 Mar 09;293(10):1223-1238. [doi: 10.1001/jama.293.10.1223] [Medline: 15755945]
60. Mastrianni A, Sarcevic A, Hu A, Almengor L, Tempel P, Gao S, et al. Transitioning cognitive aids into decision support platforms: requirements and design guidelines. ACM Trans Comput Hum Interact 2023 Jun 10;30(3):1-28. [doi: 10.1145/3582431] [Medline: 37694216]
61. Sibbald M, Monteiro S, Sherbino J, LoGiudice A, Friedman C, Norman G. Should electronic differential diagnosis support be used early or late in the diagnostic process? A multicentre experimental study of Isabel. BMJ Qual Saf 2022 Jun 05;31(6):426-433 [FREE Full text] [doi: 10.1136/bmjqs-2021-013493] [Medline: 34611040]
62. Graber ML, Mathew A. Performance of a web-based clinical diagnosis support system for internists. J Gen Intern Med 2008 Jan 19;23(Suppl 1):37-40 [FREE Full text] [doi: 10.1007/s11606-007-0271-8] [Medline: 18095042]
63. Dr. Mayson. Humei. URL: http://www.huimei.com/specialist [accessed 2024-01-30]
64. Camacho J, Zanoletti-Mannello M, Landis-Lewis Z, Kane-Gill SL, Boyce RD. A conceptual framework to study the implementation of clinical decision support systems (BEAR): literature review and concept mapping. J Med Internet Res 2020 Aug 06;22(8):e18388 [FREE Full text] [doi: 10.2196/18388] [Medline: 32759098]
65. Sutton RT, Pincock D, Baumgart DC, Sadowski DC, Fedorak RN, Kroeker KI. An overview of clinical decision support systems: benefits, risks, and strategies for success. NPJ Digit Med 2020 Feb 06;3(1):17. [doi: 10.1038/s41746-020-0221-y] [Medline: 32047862]
66. Roshanov PS, Fernandes N, Wilczynski JM, Hemens BJ, You JJ, Handler SM, et al. Features of effective computerised clinical decision support systems: meta-regression of 162 randomised trials. BMJ 2013 Feb 14;346(feb14 1):f657-f657 [FREE Full text] [doi: 10.1136/bmj.f657] [Medline: 23412440]
67. Campbell R. The five "rights" of clinical decision support. J AHIMA 2013 Oct;84(10):42-7; quiz 48. [Medline: 24245088]
68. Harrold LR, Field TS, Gurwitz JH. Knowledge, patterns of care, and outcomes of care for generalists and specialists. J Gen Intern Med 1999 Aug;14(8):499-511 [FREE Full text] [doi: 10.1046/j.1525-1497.1999.08168.x] [Medline: 10491236]
69. Fuster V, Hirshfeld JW, Brown AS, Brundage BH, Fye W, Lewis RP, et al. Working group 8: Defining the different types of cardiovascular specialists and developing a new model for training general clinical cardiologists. J Am Coll Cardiol 2004 Jul 21;44(2):267-271 [FREE Full text] [doi: 10.1016/j.jacc.2004.05.027] [Medline: 15261917]
70. Paydar P, Ebrahimpour S, Zehtab Hashemi H, Mohamadi M, Namazi S. Design, development and evaluation of an application based on clinical decision support systems (CDSS) for over-the-counter (OTC) therapy: an educational interventions in community pharmacists. J Adv Med Educ Prof 2023 Apr;11(2):95-104 [FREE Full text] [doi: 10.30476/JAMP.2022.95843.1661] [Medline: 37113680]
71. Shah C, Davtyan K, Nasrallah I, Bryan RN, Mohan S. Artificial intelligence-powered clinical decision support and simulation platform for radiology trainee education. J Digit Imaging 2023 Feb 24;36(1):11-16 [FREE Full text] [doi: 10.1007/s10278-022-00713-9] [Medline: 36279026]
72. Litvin CB, Davis KS, Moran WP, Iverson PJ, Zhao Y, Zapka J. The use of clinical decision-support tools to facilitate geriatric education. J Am Geriatr Soc 2012 Jun;60(6):1145-1149. [doi: 10.1111/j.1532-5415.2012.03960.x] [Medline: 22642270]

XSL•FO
RenderX

73. Strauss AT, James TW, Mathews SC. Fellow education improved through mobile clinical decision support application: a multi-center approach involving peri-procedural antithrombotic use. Gastroenterology 2018 Dec;155(6):2014-2015. [doi: 10.1053/j.gastro.2018.08.034] [Medline: 30165048]
74. Gong Y, Kang H. Usability and clinical decision support. In: Berner ES, editor. Clinical Decision Support Systems: Theory and Practice. Cham: Springer International Publishing; 2016:69-86.

Abbreviations

AHRA: Asian Heart Rhythm Association
ARVC: arrhythmogenic right ventricular cardiomyopathy
CAD: coronary artery disease
CDS: clinical decision support
CDSS: clinical decision support system
DDSS: diagnostic decision support system
ECG: electrocardiogram
ESC: European Society of Cardiology
SCD: sudden cardiac death
VT: ventricular tachycardia

Edited by A Kushniruk; submitted 25.12.23; peer-reviewed by H Younes; comments to author 25.01.24; revised version received 15.02.24; accepted 02.03.24; published 26.03.24.

Please cite as:
Hu Z, Wang M, Zheng S, Xu X, Zhang Z, Ge Q, Li J, Yao Y
Clinical Decision Support Requirements for Ventricular Tachycardia Diagnosis Within the Frameworks of Knowledge and Practice: Survey Study
JMIR Hum Factors 2024;11:e55802
URL: https://humanfactors.jmir.org/2024/1/e55802
doi:10.2196/55802
PMID:38530337

©Zhao Hu, Min Wang, Si Zheng, Xiaowei Xu, Zhuxin Zhang, Qiaoyue Ge, Jiao Li, Yan Yao. Originally published in JMIR Human Factors (https://humanfactors.jmir.org), 26.03.2024. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on https://humanfactors.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Usability of an Automated System for Real-Time Monitoring of Shared Decision-Making for Surgery: Mixed Methods Evaluation

Christin Hoffmann[1], BA, MSc, PhD; Kerry Avery[1], PhD; Rhiannon Macefield[1], PhD; Tadeáš Dvořák[1], MSci; Val Snelgrove[2], HNC; Jane Blazeby[1], MD; Della Hopkins[3], PhD; Shireen Hickey[4], MBChB; Ben Gibbison[1,5], MD; Leila Rooshenas[1], PhD; Adam Williams[3], MBChB; Jonathan Aning[3], BMedSci; Hilary L Bekker[6,7], PhD; Angus GK McNair[1,3], PhD; The ALPACA Study Team[3]

[1]National Institute for Health and Care Research Bristol Biomedical Research Centre, Bristol Centre for Surgical Research, Bristol Medical School: Population Health Sciences, University of Bristol, Bristol, United Kingdom

[2]Patient representative, Bristol, United Kingdom

[3]North Bristol NHS Trust, Bristol, United Kingdom

[4]Improvement Academy, Bradford Royal Infirmary, Bradford Teaching Hospitals NHS Foundation Trust, Bradford, United Kingdom

[5]University Hospitals Bristol and Weston NHS Foundation Trust, Bristol, United Kingdom

[6]Leeds Unit of Complex Intervention Development (LUCID), Leeds Institute of Health Sciences, School of Medicine, University of Leeds, Leeds, United Kingdom

[7]The Research Centre for Patient Involvement (ResCenPI), Department of Public Health, Aarhus University, Central Denmark Region, Denmark

Corresponding Author:
Christin Hoffmann, BA, MSc, PhD
National Institute for Health and Care Research Bristol Biomedical Research Centre, Bristol Centre for Surgical Research
Bristol Medical School: Population Health Sciences
University of Bristol
Canynge Hall
39 Whatley Road
Bristol, BS8 2PS
United Kingdom
Phone: 44 117 455 5993
Email: c.hoffmann@bristol.ac.uk

Abstract

Background: Improving shared decision-making (SDM) for patients has become a health policy priority in many countries. Achieving high-quality SDM is particularly important for approximately 313 million surgical treatment decisions patients make globally every year. Large-scale monitoring of surgical patients' experience of SDM in real time is needed to identify the failings of SDM before surgery is performed. We developed a novel approach to automating real-time data collection using an electronic measurement system to address this. Examining usability will facilitate its optimization and wider implementation to inform interventions aimed at improving SDM.

Objective: This study examined the usability of an electronic real-time measurement system to monitor surgical patients' experience of SDM. We aimed to evaluate the metrics and indicators relevant to system effectiveness, system efficiency, and user satisfaction.

Methods: We performed a mixed methods usability evaluation using multiple participant cohorts. The measurement system was implemented in a large UK hospital to measure patients' experience of SDM electronically before surgery using 2 validated measures (CollaboRATE and SDM-Q-9). Quantitative data (collected between April 1 and December 31, 2021) provided measurement system metrics to assess system effectiveness and efficiency. We included adult patients booked for urgent and elective surgery across 7 specialties and excluded patients without the capacity to consent for medical procedures, those without access to an internet-enabled device, and those undergoing emergency or endoscopic procedures. Additional groups of service users (group 1: public members who had not engaged with the system; group 2: a subset of patients who completed the measurement system) completed user-testing sessions and semistructured interviews to assess system effectiveness and user satisfaction. We conducted quantitative data analysis using descriptive statistics and calculated the task completion rate and survey response rate

XSL•FO
RenderX

(system effectiveness) as well as the task completion time, task efficiency, and relative efficiency (system efficiency). Qualitative thematic analysis identified indicators of and barriers to good usability (user satisfaction).

Results: A total of 2254 completed surveys were returned to the measurement system. A total of 25 service users (group 1: n=9; group 2: n=16) participated in user-testing sessions and interviews. The task completion rate was high (169/171, 98.8%) and the survey response rate was good (2254/5794, 38.9%). The median task completion time was 3 (IQR 2-13) minutes, suggesting good system efficiency and effectiveness. The qualitative findings emphasized good user satisfaction. The identified themes suggested that the measurement system is acceptable, easy to use, and easy to access. Service users identified potential barriers and solutions to acceptability and ease of access.

Conclusions: A mixed methods evaluation of an electronic measurement system for automated, real-time monitoring of patients' experience of SDM showed that usability among patients was high. Future pilot work will optimize the system for wider implementation to ultimately inform intervention development to improve SDM.

International Registered Report Identifier (IRRID): RR2-10.1136/bmjopen-2023-079155

(JMIR Hum Factors 2024;11:e46698) doi:10.2196/46698

KEYWORDS

surgery; shared decision-making; patient participation; mixed methods; surgery; real-time measurement; patient-reported measure; electronic data collection; usability; data collection; patient reported; satisfaction; mobile phone

Introduction

Background

Contemporary health care puts patient-centered care at the heart of its delivery [1-4]. Shared decision-making (SDM) is a form of communication that promotes a dialogue between those involved in making health care choices. Therefore, treatment decisions are based on a shared understanding between patients and health care professionals of the evidence base for treatment and prognosis, patient values, preferences and beliefs, and clinical reasoning to personalize service delivery [5]. SDM is desired by patients and has become a key priority for health care systems globally [6-9]. Ensuring high-quality SDM when discussing and deciding treatments with patients can have many benefits, such as reduced information asymmetry or health service use [10,11]. It has been shown to contribute to good patient outcomes and satisfaction [12-15].

Globally, approximately 310 million operations are performed annually [16]. Surgery is often the only available treatment for a wide variety of minor and major medical conditions, and people increasingly choose surgical treatment (5.3% increase from 2009 to 2014 in the United Kingdom) [17]. Improving surgical patients' experience of SDM before surgery is particularly important because the effects of surgery are immediate and nonreversible. Patients cannot decide to discontinue treatment if the benefits fall short of expectations or side effects become unacceptable. Furthermore, making good surgical decisions may avoid negative impacts on health service costs (eg, through canceled operations) and patient outcomes [18-20].

Strategies aimed at improving SDM in complex health care settings can range from communication skills workshops for health care professionals [21] to educational videos [22] and booklets for patients [23]. However, their effects are mixed [14,15]. Systematic reviews of evidence to improve SDM conclude that achieving long-term change is likely to necessitate interventions that support the implementation of strategies at the organization, clinician, and patient levels [24-26]. However, there is uncertainty about how to realize change on a large scale across health care systems [27-33]. One recommended way to achieve this is through routine monitoring of patients' experience of SDM [34], but robust methods are lacking. Existing approaches to data collection are delayed, potentially affecting patients' accounts of their experience and impacting the ability to respond quickly and effectively before surgical treatments. Advances in technology mean that novel approaches to assessing patients' experiences of SDM can incorporate automated, electronic data capture close to the point of treatment consultations. This offers opportunities for providing information more accurately and in a timely manner, offering an effective way to develop interventions to improve SDM before surgery. Systems routinely collecting electronic patient-reported measure (ePRM) data in other contexts have been shown to improve care and outcomes for patients, including quality of life outcomes in pediatric dermatology [35] or symptom reporting in chronic kidney disease [36]. We developed a novel system to routinely monitor patients' experience of SDM automatically and in real time.

The evaluation of existing ePRM systems highlights the importance of user-friendly processes for their optimal performance [37-40]. Furthermore, the principles of good usability are important because they can be vital to the widespread uptake of ePRM systems by patients and their successful implementation in clinical practice [41-43]. Usability is an outcome defined as the extent to which the system can be used by specified users [44]. Several methods are available to evaluate measures of usability in health care [45-49]. A widely used framework contains standards set by the International Organization for Standardization (ISO) [50,51]. The guidelines recommend evaluating and optimizing the concepts of system effectiveness (the ability of participants to complete the survey), system efficiency (resources required to complete the questionnaire), and user satisfaction (subjective opinions of participants' experience with the measurement system) to achieve good usability.

XSL•FO
RenderX

Aim and Objectives

We aimed to examine the usability of a novel, automated, real-time measurement system to monitor surgical patients' experience of SDM. The specific objectives were to evaluate the measurement system's (1) effectiveness, (2) efficiency, and (3) user satisfaction among a large sample of surgical patients from a wide range of surgical specialties.

Methods

We used quantitative and qualitative methods to examine usability by evaluating the indicators and metrics related to system effectiveness, system efficiency, and user satisfaction. This study adhered to the ISO guideline 9241-11:2018 and followed recommendations for the usability testing of electronic patient-reported outcome measures [49,51].

Context and Setting

This study is part of a wider project to develop, pilot, and evaluate a decision-support intervention that uses real-time monitoring of patients' experiences to improve SDM (the ALPACA Study [52]). The project was initially set up as a quality improvement project at a large acute National Health Service (NHS) Trust in England, United Kingdom, which provides a range of acute and specialized clinical care services in South West England.

To facilitate automated, real-time data collection of patients' experience of SDM, a customizable off-the-shelf ePRM system (Cemplicity) was procured from a third-party software provider in March 2021. The software provider is an ISO 2001 certified, NHS-authorized ePRM provider, compliant with necessary accessibility and health data governance standards (eg, General Data Protection Regulations and Digital Technology Assessment Criteria). Before deployment and customization, the software provider tested the system development and design. Specifically, the prior rollout of the software across 6 countries and over 3000 health care institutions incorporated feedback from users across different health care settings and patients of diverse age groups, technology literacy, and health confidence. All measurement system interfaces are mobile optimized.

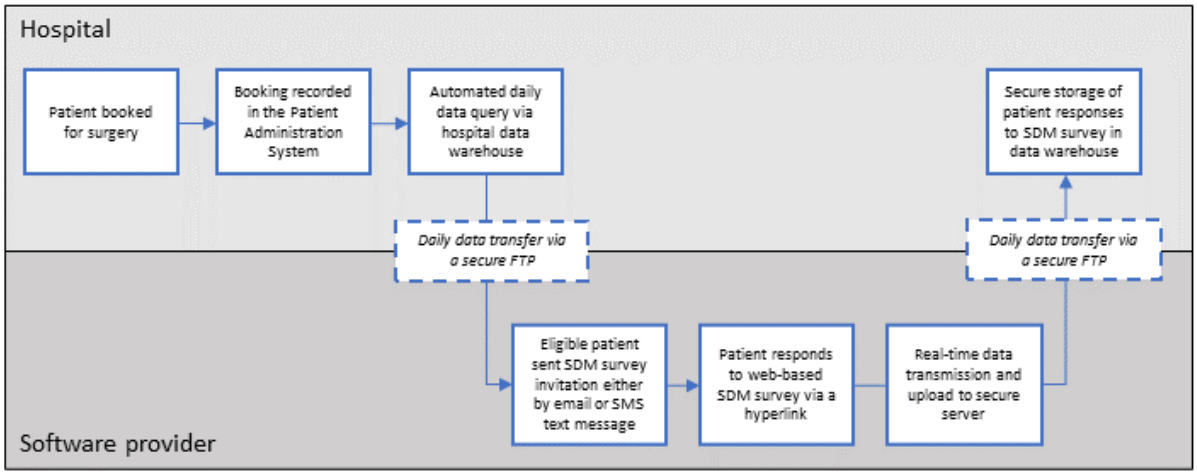
Customization for the purpose of this study was undertaken in collaboration with the software provider and included adapting the following: (1) the system's content and layout to include instruments to assess patients' experience of SDM and (2) data capture mechanisms to implement the system in the NHS Trust.

To assess patients' experience of SDM, 2 validated and widely used patient-reported measures were selected to measure SDM (CollaboRATE and SDM-Q-9). These were chosen by consensus within the study team, which was informed by a systematic review of SDM measurement instruments [53], national guidelines [25], and recommendations and use within the NHS clinical practice [34,54,55]. CollaboRATE is a 3-item instrument measured on a 10-point scale with answer options ranging from 0 ("no effort was made") to 9 ("every effort was made"). SDM-Q-9 consists of 9 items measured on a 6-point scale with answer options ranging from "completely disagree" to "completely agree." The measurement properties of both instruments have been demonstrated to be acceptable [56,57]. The measurement instruments were operationalized into a 12-question electronic survey format, branded to match the NHS Trust guidelines, and integrated into the patient-facing measurement system. Screenshots of the customized content are presented in Multimedia Appendix 1.

To implement the measurement system, secure data exchange processes were established between the software provider and the NHS Trust's information technology system and subsequently widened to various patient cohorts within the surgical departments. Specifically, SQL data queries were developed to identify and extract details of patients booked for surgery from the electronic patient record system that routinely records the patients' demographic and clinical information. The queries were designed to run automatically, securely transferring data from the hospital to the software provider on a daily basis. The 2 SDM measures were administered to patients upon being booked for surgery, with invitations sent either by email or SMS text messaging if no email address was available. Patient responses were received and processed using the measurement system. A reciprocal data feed securely returned response data to the hospital data warehouse for secure storage. A flow diagram of the measurement system process is provided in Figure 1.

XSL•FO
RenderX

Figure 1. Flow diagram of the process of automated real-time shared decision-making (SDM) monitoring through the measurement system. FTP: file transfer protocol.



Study Steering Group

A multidisciplinary study steering group was convened and consisted of a patient and public contributor, health care professionals, methodologists, social scientists, statisticians, and health services researchers. Regular meetings ensured the group's strategic oversight throughout and sought their input into the study design, research activities, and analyzing and interpreting results.

Patient and Public Involvement

We invited a patient and public contributor with lived experience of surgery to the study steering group, which was set up as part of the wider project. The input was sought from the patient and public contributor as appropriate throughout the study (eg, review of patient-facing materials, including survey invitation and instructions, and interim findings from qualitative analyses). In addition, we organized a patient and public advisory meeting which 6 public contributors attended for 1 hour via a Zoom (Zoom Video Communications, Inc) meeting. The aim of the meeting was to obtain patient and public perspectives on the overall project plan and its key challenges. The topics discussed included recruitment, acceptability, and satisfaction with the measurement system, which informed the design aspects of this study.

Usability Concepts

The usability of the measurement system was examined by evaluating metrics and indicators relevant to 3 concepts, including system effectiveness, system efficiency, and user satisfaction. The definitions are summarized in Textbox 1, and the details of their assessment are described subsequently.

Textbox 1. Definitions of usability concepts.

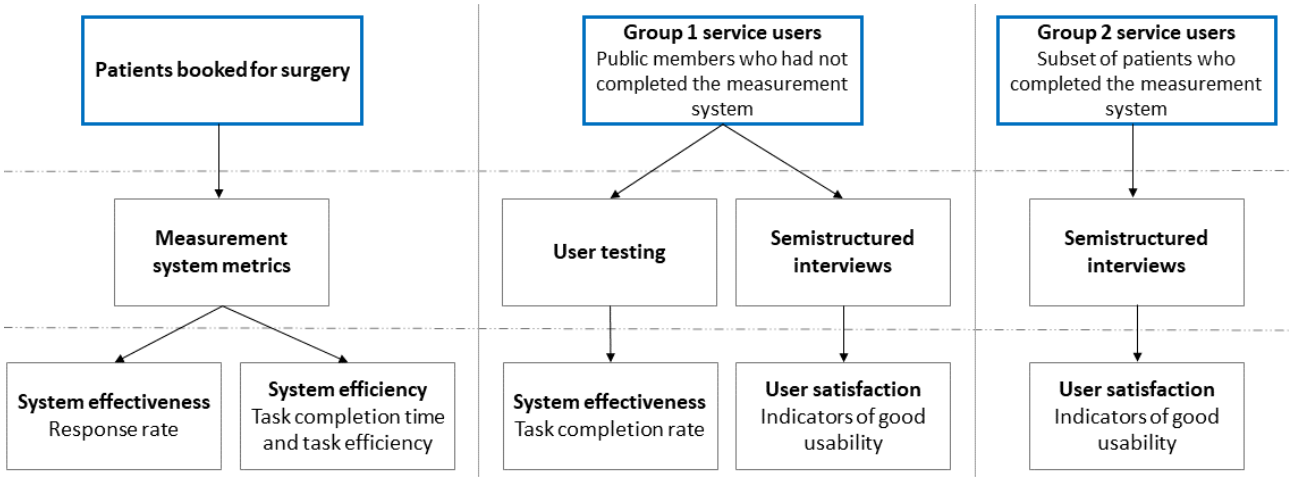
- System effectiveness: the ability of participants to perform tasks to achieve predetermined goals completely and accurately, without negative consequences (eg, poor layout of the system interface leading to participants missing or accidentally selecting system options) [36,49-51].
- System efficiency: the amount of participant resources required to achieve the prespecified goals [49,58].
- User satisfaction: the subjective opinions of the participants based on their experience of interacting with the system [49]. This includes any subjective reports about likes, dislikes, and recommendations for changes [51].

Participants and Procedures

We used multiple cohorts of participants and procedures for quantitative and qualitative data collection for this study. Figure 2 illustrates the different cohorts of participants and provides an overview of the data collection procedures used to evaluate the usability concepts.

XSL•FO
RenderX

Figure 2. Overview of participant cohorts, data collection procedures, and usability concepts.



Participants and Recruitment

To obtain quantitative measurement system metrics to assess system effectiveness and efficiency (refer to the Quantitative Analysis section for further details), automated, real-time data collection was conducted between April 1 and December 31, 2021, and rolled out across 7 surgical departments: orthopedic, urology, gynecology, neurosurgery, gastrointestinal, vascular and breast. We included adult patients booked for elective surgery in these 7 specialties. Patients aged <18 years, those without the capacity to consent for medical procedures, those undergoing emergency and endoscopic procedures, and those without access to an appropriate internet-enabled device (ie, mobile phone, smartphone, PC, tablet, or similar device) were excluded.

We recruited 2 further groups of service users for user testing and interviews to obtain quantitative and qualitative data to assess system effectiveness and user satisfaction.

Group 1 participants were individuals who had not engaged with the measurement system before user-testing sessions to ensure naive user interactions [59] (refer to the User Testing section for detailed user-testing methods). Service users with experience of surgery were recruited through patient experience panels within 2 NHS Trusts (North Bristol NHS Trust and Bradford Teaching Hospitals NHS Foundation Trust). A panel coordinator identified and approached potential participants via email containing a recruitment advertisement. Sampling was purposive to achieve the maximum possible variation in recognized protected characteristics (eg, sex, disability, and race) and experience of surgery.

Group 2 participants were individuals who had engaged with the measurement system to explore user satisfaction after interacting with the system (refer to the *Semistructured Interviews* section for more details). These were a subset of eligible patients who completed the measurement system. A member of staff with authorized access to the patient administration system and patient response data stored in the data warehouse recruited participants via telephone. We used a purposive sampling strategy to achieve variation in characteristics, including age, ethnicity, sex, type of surgery received, and experience of good or bad SDM (identified through survey responses).

Procedures

Measurement System Metrics

Relevant metrics automatically collected by the measurement system were used to examine usability quantitatively (eg, responses to questionnaire items and timestamps for starting and submitting the survey). Unique entries were recorded for each patient who received the invitation to complete the measurement system. Entries and corresponding data collected between April 1 and December 31, 2021, were available for analysis.

User Testing

Postdeployment user-testing sessions were conducted between June and December 2021 and were performed in a simulated environment.

Group 1 participants were invited to participate in a one-to-one 1-hour videoconference via Zoom with a researcher to complete the measurement system live. Sessions began by reminding participants about the aim and the process of user testing the measurement system. Service users were then sent an SMS text message or email invitation (depending on their preference) that included a test link to the survey. Specific user-testing links were set up to allow simulated completion of the measurement system (ie, responses were not used for live response data). Sessions assessed system effectiveness (including any issues related to system functionality or completion). A concurrent think-aloud technique was applied to vocalize reactions and thinking processes [60-62], supplemented with observational notes of any difficulties encountered [63,64].

User-testing sessions were conducted by 1 member of the study team who had experience in think-aloud methods (AGKM or CH). A topic guide was developed to guide conversations (Multimedia Appendix 2). Sessions were audio recorded and transcribed using unique identifiers to ensure anonymity. Field notes and any problems during the measurement system completion were recorded in a table using Excel (Microsoft Corp).

XSL•FO
RenderX

Semistructured Interviews

We conducted semistructured, in-depth interviews using retrospective probing to explore the service users' views about the indicators of usability of the measurement system [65].

Interviews were conducted with group 1 participants following user testing via the same web-based videoconferencing software. Group 2 participants were invited to take part in an approximately 30- to 45-minute phone or videoconferencing call (according to their preference) during which they reflected on completing the measurement system. The conversations followed a previously tested and refined topic guide that was based on standard usability concepts [51]. An example topic guide can be found in Multimedia Appendix 2.

Interviews were performed by either of the 2 researchers (AGKM or CH), audio recorded, and anonymized during transcription.

Analysis

Quantitative Analyses

All quantitative analyses were performed by 3 researchers (TD, AGKM, and CH) using the statistical software package STATA (version 16.0; StataCorp LLC).

System Effectiveness

We assessed system effectiveness by calculating the user task completion rate based on usability testing sessions and the survey response rate based on measurement system metrics [66].

The user task completion rates were calculated as a percentage of tasks completed by the total number of tasks. A process map was created defining the number and type of tasks (or steps) required to complete the measurement system. Successful completion means that all tasks were completed without user errors. User errors were deviations or problems encountered that interfered with successful task completion. Noncritical errors were defined as those that were successfully addressed by the testers themselves following instructions from the observer. Critical errors were those that required the observer to intervene or take remedial actions.

The survey response rate was calculated as a percentage (number of completed surveys/number of patients invited × 100). Surveys were considered complete when responses to all 3 items of the CollaboRATE measure and at least 7 out of 9 items of the SDM-Q-9 measure were returned.

System Efficiency

We assessed system efficiency by calculating the task completion time and task efficiency based on measurement system metrics [58,66].

The task completion time was defined as the time participants took from the first activity (starting the survey by following the hyperlink) to the last activity (submission of the survey). Task efficiency was defined as the time spent to complete each task (timestamps were recorded in the following format: hh:mm:ss). Analyses were based on those who completed the measurement system for whom typical first and last activity timestamps were available (ie, atypical timestamps were those with no recorded activity time). Extreme outliers were excluded because the system allowed service users to leave and later return to the survey and continue submission (eg, the next day or the following week). These were defined as those with the task completion time >3 times the IQR [67].

Qualitative Analyses of User Satisfaction

User satisfaction was assessed by evaluating service users' self-reported experiences of using the system through user-testing sessions and semistructured interviews [66]. Discussions explored perceptions of usability aspects, including service users' interpretation of the system's ease of use and navigation, their satisfaction with instructions and visual display, and the likelihood of using the system again or recommending it to others.

Transcripts obtained from user-testing sessions and semistructured interviews were uploaded to the qualitative data management software NVivo (version 20.5.1; QRS International) and analyzed using thematic analysis [68]. This involved systematic coding of data to identify commonly mentioned concepts within the data and to develop themes and subthemes. The coding was inductive and iterative and followed predefined steps of data familiarization, generation of initial codes, and searching for themes. Coding was performed by 2 researchers independently who met regularly to review themes. Analysis and interpretation of qualitative data were further supported in 2 ways. First, a report of the interim findings was produced and discussed with the wider multidisciplinary steering group. Second, the presentation of interim findings to the patient and public advisory group sought further input.

Ethical Considerations

This study was part of a project spanning quality improvement and research. Therefore, it was subject to 2 governance processes requiring separate approvals. Monitoring patients' experience of SDM in routine clinical practice was initially approved through a quality improvement proposal at North Bristol NHS Trust (reference: Q80008). This was then incorporated into a larger program of work, where all processes were approved through the appropriate governance framework (Consent and SDM Program Board, reporting to the Clinical Effectiveness and Audit Committee). Ethics approval for conducting interviews with NHS patients was granted by the NHS Health Research Authority North West – Liverpool Central Research Ethics Committee (reference: 21/PR/0345). Participants provided electronic consent through a link to a secure data management platform (version 11.1.18, REDCap [Research Electronic Data Capture]; Vanderbilt University) [69] before any study activity commenced.

Results

Participants and Procedures

A total of 5794 surgical patients received invitations to complete the survey and for whom unique entries were recorded in the measurement system. Of these, 2254 returned the completed surveys (refer to Table 1 for patient characteristics) and provided data for the analysis of measurement metrics.

XSL•FO
RenderX

Table 1. Characteristics of patients who completed the measurement system (N=2254).

Characteristics	Patients, n (%)
Sex	
Female	1243 (55.15)
Male	1011 (44.85)
Age group (y)	
<29	170 (7.54)
30 to 39	213 (9.45)
40 to 49	277 (12.29)
50 to 59	529 (23.47)
60 to 69	555 (24.62)
70 to 79	402 (17.83)
≥80	108 (4.79)
Ethnicity[a]	
Other	104 (4.61)
White British	977 (43.35)
Specialty	
Breast	278 (12.33)
Colorectal	67 (2.97)
General	194 (8.61)
Gynecology	106 (4.7)
Neuro	288 (12.78)
Trauma, orthopedics and spinal	555 (24.62)
Upper gastrointestinal	41 (1.82)
Urology	584 (25.91)
Vascular	141 (6.26)

[a]Missing data: n=1173.

A total of 25 service users (group 1: n=9; group 2: n=16) participated in user-testing sessions and semistructured interviews.

In group 1, a total of 9 service users completed 8 user-testing sessions. Most sessions were completed on a one-to-one basis (7/9, 78%). One session was completed with 2 participants, which included 1 service user with disability and their caregiver who provided additional support. All sessions were held via videoconference and lasted for an average duration of 43 (SD 15.1; range 29-78) minutes. Service users in this group were mostly female participants (6/9, 67%) and self-identified as Asian (1/9, 11%), other White background (1/9, 11%), and White British (7/9, 78%). Details about the surgical experience were known for 4 service users who represented orthopedic (2/4, 50%), upper gastrointestinal (1/4, 25%), and ophthalmic (1/4, 25%) specialties.

In group 2, 16 service users completed semistructured interviews between June and November 2021. Most interviews were conducted via telephone (15/16, 94%), with 1 (6%) interview conducted via videoconference, lasting for an average duration of 36 (SD 9.9; range 21-50) minutes. Most service users in group 2 were female participants (10/16, 62%) and were 51 (SD 15.8) years on average. All participants were from a White British background (16/16, 100%). Efforts were made to recruit participants from a wide range of ethnic minority backgrounds; however, due to a large amount of missing data (Table 1), this was unsuccessful. The characteristics of group 2 participants are presented in Table 2.

XSL•FO
RenderX

Table 2. Characteristics of group 2 service users (n=16).

Characteristics	Service users, n (%)
Age (y), mean (SD; range)	51 (15.8; 23-80)
Sex, n (%)	
Female	10 (62)
Male	6 (38)
Ethnicity, n (%)	
White British	16 (100)
Surgery type, n (%)	
Breast	3 (19)
Colorectal	2 (13)
General	2 (13)
Gynecology	1 (6)
Trauma and orthopedics	2 (13)
Urology	5 (31)
Vascular	1 (6)

Usability Concepts

System Effectiveness

A process map to assess task completion contained 19 tasks (or steps) required to complete the measurement system. Tasks ranged from "Open text message/email" to "Click on 'Submit'" and are detailed in Multimedia Appendix 3.

A total of 171 tasks across 8 user-testing sessions were submitted by all 9 group 1 participants. One service user reported 2 noncritical errors across 2 tasks when completing the measurement system using a mobile phone. The first error occurred following task 1 "Open text message." This forced an additional step to resolve a pop-up notification which prompted the service user to select an internet browser to open the survey link. The second error occurred following task 5 "Select response to question 1." The displayed answer options for CollaboRATE item 1 were cut off at 8, not presenting answer option 9 (every effort was made). Further scrolling was required by the service user to be able to select the answer option 9. Both noncritical errors were managed and resolved without requiring observer input. Consequently, a total completion rate of 98.8% (169/171) was achieved. No critical errors or failures in completing the tasks were reported.

The survey response rate was 38.9% (2254 completed surveys/5794 patients invited × 100).

System Efficiency

Out of the 2254 responses available, 1106 (49.07%) were excluded from analysis. These 1106 responses included 719 (65.01%) responses with an atypical timestamp (ie, no activity time was recorded because the timestamp for the first and last activity was 00:00:00, which was identified as a technical issue and rectified by the software provider) and 387 (34.99%) responses identified as extreme outliers (ie, the task completion time was >12 min). Assessment of the completion time of 1148 (50.93%) of the 2254 responses showed that service users required an average median duration of 3 (IQR 2-4) minutes to complete the measurement system. Calculations of task efficiency showed that the average median time taken per task was 9 (IQR 6-13) seconds.

User Satisfaction

Analysis of qualitative data from user-testing sessions and semistructured interviews with a subset of patients revealed four main themes related to user satisfaction as follows: (1) acceptability, (2) ease of access to the system, (3) ease of use, and (4) satisfaction with the measurement system.

Acceptability

Indicators of Good Acceptability

Service users who were interviewed as part of the qualitative data collection frequently commented on the low burden of completing the measurement system, suggesting good acceptability among the participants. This was mainly because of the low number of questions contributing to the measurement system being considered quick and straightforward to use:

> *Short survey, key thing—not too much of your time.* [PT9, group 1]
>
> *I did it from my phone so yes it was very straightforward.* [PT13, group 2]
>
> *I don't remember feeling any burden [...], it was quite easy.* [PT19, group 2]
>
> *I don't think it seemed too long. It was enough. To be honest, if it had been a lot more, I probably wouldn't bother to do it.* [PT21, group 2]

Furthermore, service users highlighted the common use of web-based surveys to obtain feedback in health care and other general settings. Therefore, they felt a certain level of familiarity with the measurement system, which contributed to the good acceptability:

XSL•FO
RenderX

> *I thought it was, I mean, pretty standard, you know, arial buttons, nought to ten on how much you disagree, agree, disagree to something so yes, familiar with many other surveys that I've seen before.* [PT12, group 2]

Potential Barriers to Acceptability

Some barriers to completing the measurement system were highlighted. For example, participants mentioned that the service users may easily ignore or forget to complete the measurement system as follows:

> *It's easy not to [complete the measurement system], I've had them from places, not about health or anything important like that, but it's easy just [to] think, "Oh, I'll do that later," and then never go back to it.* [PT7, group 1]

Another example included concerns about the number of SMS text messages and surveys received from other sources and the cumulative burden:

> *I mean the good thing about it is it's simple and easy and you just get the nudge, but on the other hand there are lots of other nudges coming through at you.* [PT19, group 2]

This contributed to a small number of service users questioning the credibility of the invitation to complete the measurement system:

> *Something came through via email which to be honest I wasn't sure if it was a genuine thing or if it was something else.* [PT10, group 2]

Solutions to Address Barriers

Service users were asked about the usability of solutions to address these issues and included support for reminder emails:

> *One follow-up is a good idea but not more than one possibly because then people start to feel a bit harassed, but I think a second one is a good idea because of the forgetting thing and they go oh yeah, I'll do it this time.* [PT4, group 1]

Service users thought that the use of email would address this problem for some service users:

> *It is at the top of my email pile again, I'd better do it, so it jogs your memory, texts don't do that, it's a very momentary thing, text messaging.* [PT13, group 2]

Furthermore, service users suggested to increase the personal relevance and awareness of the measurement system:

> *You've got to feel that you're going to benefit, and it's really relevant to you, for you to have the interest to do it.* [PT7, group 1]
>
> *They [service users] really, really need to know it's coming because I don't know about you but we're very, very careful what we open and if this just appeared with no warning I wouldn't open it.* [PT4, group 1]

Service users mentioned the need to highlight the brevity of the measurement system and the low number of questions:

> *There are people who will fill them in if they're told it's very short, which is why it's important that it says it's short.* [PT9, group 1]
>
> *I think sometimes if you open one you can see that it's 100 questions you just think I probably won't do that.* [PT14, group 2]

Ease of Access to the System

Indicators of Good Ease of Access to the System

All service users were able to access the measurement system without problems and commented on its ease of access through both methods, email and SMS text message:

> *I think most people nowadays are comfortable with computers and technology.* [PT14, group 2]

Some service users expressed a preference for using either email or their phone to complete the measurement system. However, there was no conclusive evidence to suggest the superiority of either email or SMS text message:

> *I guess that for me making it [come to my phone] makes it more accessible 'cos you don't have to go in your emails. It automatically comes through and you can do it at any time and reply at any time, so you can do it when it's convenient to you and its literally just a text on your phone.* [PT1, group 1]
>
> *Although I use a smart phone quite a lot, sometimes it's difficult to manipulate it, whilst a laptop I find much more easier to use.* [PT8, group 1]

Furthermore, service users commented on the good comprehensibility and legibility of the content, contributing to good levels of ease of access to the system. For example, comments included that there was a sufficiently large font option for those who required or preferred larger screens:

> *I think the presentation of it on my phone, and I don't have a large phone, I just have a small phone, I could read all that quite easily.* [PT7, group 1]
>
> *They were really easy to understand[...] The questions were very clear, I thought they were quite well[...] focused and well explained.* [PT24, group 2]

Potential Barriers to Access to the System

Some service users expressed concerns regarding the system's ease of access for certain population groups. Most frequently, concerns were raised in connection with older adults and lack of access to technology. Furthermore, considerations included the ease of access to the measurement system for non–English-speaking service users and those with disabilities:

> *There's also a certain cohort would be using online. [...] I do think people will miss out but if it's just being pinged… whether it's on text or email* [PT3, group 1]
>
> *People that English isn't their first language, that could be a bit of a consideration.* [PT5, group 1]

XSL•FO
RenderX

Solutions to Address Barriers

The most frequently mentioned solutions were common alternatives to electronic data collection in connection with support measures for questionnaire completion:

> *I mean there's probably still a gap with the older generation who wouldn't be comfortable doing it, and would prefer doing it via communication of phone or in written format.* [PT14, group 2]

Ease of Use

Most often, the simplicity of the system was highlighted in connection with the ease of completing the measurement system. Furthermore, the ease of use was often attributed to the brevity of the measurement system:

> *I actually thought it was quite simple and quite straightforward and easy.* [PT3, group 1]
>
> *That has been perfectly straightforward, for someone who's not very IT literate, that was all fine.* [PT7, group 1]
>
> *Yeah, that was very easy, it didn't take very long [...] I remember it did seem simple* [PT23, group 2]

Moreover, most service users commented on the visual display, which was perceived as appealing and very clear. The clear layout of the survey contributed to high comprehensibility among participants:

> *It is very clear and also I quite like the bold type. [...] very clear again and very easy to read.* [PT3, group 1]
>
> *It's pretty obvious straight off of that where the survey has come from including the logo and almost like the colours of the survey match with the NHS logo [...] I think that part of it makes it really easy.* [PT21, group 2]
>
> *Yeah, that's laid out really spaced out and easy to read.* [PT6, group 1]

Service users frequently mentioned the ease of navigation and thought it was "basic and straightforward" (PT20, group 2). Others mentioned further details regarding what they liked about the navigation:

> *There is no need to zoom in or zoom out or move around a page or click buttons to find the survey so I think all of that aspect is really easy. [...] It's easy to use and the agree or disagree buttons are really straight to the point.* [PT21, group 2]

One service user also commented on the loading speed of the survey page:

> *I think it's easy to use because it doesn't take long to load which I think is important.* [PT20, group 2]

No service user raised concerns that could be considered barriers to the ease of use of the measurement system.

Overall Satisfaction With the Measurement System

All service users provided positive feedback regarding the abovementioned themes of acceptability, ease of access to the system, and ease of use, which indicated high satisfaction with the measurement system. General supportive comments were made throughout the user-testing sessions and semistructured interviews:

> *Yeah, absolutely brilliant. I'll give that 11 out of 10. [...] Somebody who designed this did a good job.* [PT5, group 1]

All respondents agreed when asked whether they are likely to complete the measurement system again:

> *Yeah, I would definitely respond to it again.* [PT6, group 1]

In addition, there were unprompted comments related to satisfaction with particular features. For example, service users pointed out that they particularly liked the "back buttons" to return to previous questions, the option to pause the measurement system and return at a different time, and the fact that there are contact details of the hospital in case this survey was received in error:

> *You've got the option, you can go back and change something, or if there was something you were worried about that you've done, it's clear that you can go back.* [PT7, group 1]

Discussion

Principal Findings

This study examined the usability of a novel automated and real-time ePRM system to monitor patients' experience of SDM in routine clinical practice. We used a large sample from a diverse range of surgical specialties to evaluate system effectiveness, system efficiency, and user satisfaction.

Overall, the evaluation of the measurement system demonstrated good usability. Metrics relevant to the effectiveness and efficiency showed that the system can be used without problems and completed quickly. The results from qualitative testing sessions and interviews with 25 service users showed that the measurement system has good user satisfaction. It was perceived as acceptable, easy to access, and easy to use. Service users identified potential barriers to acceptability and ease of access to the system, which can inform strategies for the optimization of the measurement system.

Limitations

This study has certain methodological limitations. First, we purposively selected participants to include individuals from a wide socioeconomic background with varying computer literacy skills. While this study exceeded the recommended sample size for usability testing [70-72], service users in our sample were primarily White British (23/25, 92%), English-speaking adults with capacity to consent for medical treatments, and from specific geographic areas of the United Kingdom (West, South West, and North East England). This may limit the generalizability of the study findings. It is uncertain whether the inclusion of more participants from more diverse backgrounds would have elicited different perspectives on the measurement system. Second, only patients who had completed the measurement system were eligible to participate in semistructured interviews. Data protection regulations limited

XSL•FO
RenderX

our ability to recruit individuals who had not completed the survey. Therefore, we were unable to explore whether nonengagement with the system was due to reasons related to usability not mentioned by the study participants. Barriers to engagement may align with the themes identified during semistructured interviews, which are partly addressed by ongoing work (refer to the following section). Separately, there is ongoing work which includes conducting follow-up phone calls with patients to explore the reasons for nonengagement. Third, usability may also be evaluated using validated measurement instruments to capture quantitative measures of individuals' perception of usability from a larger, representative sample size [73,74]. This study did not include such measures in addition to the ePRM to avoid distorting usability outcomes. For example, the additional length of the survey may have affected system efficiency and impacted perceptions of ease of use. Instead, we included a range of methods to assess usability to triangulate the data sources [75].

Comparison With Prior Work

Existing research has investigated optimal strategies and methods for collecting ePRMs [40,76-79]. The usability evaluation of electronic platforms is common and has been fundamental in optimizing systems to collect ePRMs across a range of health care settings [80] and also within surgery [81,82]. Less is known about systems that monitor patients' experiences automatically and in real time. We are aware of only 1 recently published protocol describing a similar measurement system [83], but we were unable to identify studies with specific relevance to surgery or SDM. Our study addresses this gap and provides insights into the usability of an automated measurement system that monitors ePRMs for SDM in real time. The measurement system in our study was evaluated for service users undergoing surgical treatment; however, the findings may be applicable to other health care settings.

Evidence of good usability of an automated measurement system that captures surgical patients' experiences in real time supports the measurement systems' potential for scalability. The use of the system is recommended in similar health care settings where policy makers or official bodies wish to audit or monitor patients' experiences of SDM or aim to inform interventions to improve SDM before treatment. System effectiveness and efficiency are central components to service users' successful interaction with any system [51]. The usability concepts evaluated have been shown to be key in other systems rolled out in surgical departments [84] and are likely to play a role in the wider adoption of the measurement system [85]. This study showed that service users were able to successfully complete the measurement system and that they required little time and effort to do so. In addition, good user satisfaction is vital to a system's sustainability and is used as a measure of the success of digital information systems within health care organizations worldwide [86-88]. User satisfaction with ePRM systems and perceived acceptability, in particular, have been shown to be key to their uptake among stakeholders [89,90]. The qualitative evidence obtained from service users in this study demonstrated good acceptability, ease of access to the system, and ease of use, which suggests low concern regarding user satisfaction. Some steps to optimize the system to address identified usability concerns and adapt SDM measurement to other care contexts [91] might be necessary before a wider rollout to other health care settings.

This study highlighted well-known barriers to ease of access to electronic measurement systems [92,93]. Specifically, literacy with electronic systems can be lower in older and frail adults and among individuals without capacity to consent [94-97]. While the measurement system response rate in this study (2254/5794, 38.9%) was notably higher compared to those reported in other studies evaluating measurement systems (eg, 18% in the study by Iversen et al [98], 20% in the study by Bliddal et al [99], or 30% in the study by Arner [82]), it may be indicative of such barriers experienced by surgical patients. The solutions to improve ease of access identified in this study include additional paper-based methods. Furthermore, barriers may be overcome through assisted data collection using a tablet computer at the point of care [100]. Additional resources may be required to ensure full and accurate data capture for adults without capacity to consent to medical treatments completing the measurement system [101]. Similarly, language barriers have been shown to affect service users' ease of access to the system and the quality of responses to ePRM systems [93,102]. Translating content can be key to addressing such language barriers, as demonstrated by widely used quality of life measures [103]. Further work is currently ongoing to address relevant issues to maximize inclusivity (ISRCTN [International Standard Randomised Controlled Trial Number] registry ID: 17951423). Specifically, this line of work seeks to explore the views of underserved groups (eg, limited income, older age, and ethnic minority groups) using qualitative methods to understand how the use of the system and future intervention development can be optimized to maximize inclusivity. This work will consider nondigital materials, translation of study materials, measurement system content, and measurement instruments using appropriate guidance [104] and will include non-English qualitative data collection. Detailed methods will be reported in a separate publication.

High-quality SDM can be a moderator and mediator of health and care quality [105], addressing the challenges of true patient-centered care (eg, reducing asymmetry in medical knowledge between patients and surgeons and addressing issues of individual preferences). To improve patients' experiences of SDM before surgery, additional intervention development work is needed to complement automated, real-time monitoring of SDM experiences. Evidence from other clinical settings suggests that interventions, including real-time feedback, in addition to routine monitoring of ePRMs, can lead to improvements in outcomes or clinical performance [81,106-108]. This study demonstrated the good usability of a measurement system that automatically collects, stores, and retrieves ePRM data and is ready to provide feedback on this information in digital format near to real time. This suggests that the system is ready to provide instantaneous feedback on surgical patients' experience of SDM to clinical teams, which has the potential to improve SDM. Future work will explore the optimal design and feasibility of feedback mechanisms and examine the acceptability of the system. Refinements to optimize the usability and inclusivity of the system are required before

XSL•FO
RenderX

evaluating the effectiveness of an intervention to improve SDM. Key to this work will be obtaining wider perspectives from other stakeholders involved in the intervention (eg, health care professionals and stakeholders from the lower-income, ethnic minority, and older age groups). In the long term, strategies to facilitate the implementation of the measurement system in routine clinical care will be investigated and evaluated using evidence-based approaches to intervention design [109].

Conclusions

We examined the usability of a measurement system for automated and real-time ePRM collection to monitor patients' experience of SDM in a large sample using 2 brief, validated instruments. The findings suggest good usability and support scalability of the measurement systems to other secondary health care institutions and will inform its optimization. Complementary work is currently exploring the feasibility and acceptability of monitoring and feedback experience of SDM with patient and professional stakeholders. Future implementation and formal evaluation of the measurement system will be performed to establish whether routine monitoring and feedback of patients' experiences has the potential to improve SDM for surgical patients.

Acknowledgments

The ALPACA Study was funded by the National Institute for Health and Care Research (NIHR) Program Development Grant (NIHR205174). Furthermore, this study was supported by an NIHR Clinician Scientist award to AM (NIHR CS-2017-17-010) and by the NIHR Biomedical Research Centre at the University Hospitals Bristol and Weston National Health Service (NHS) Foundation Trust and the University of Bristol (BRC-1215-20011, NIHR203315). TD was funded by the NIHR Academy through an NIHR Methodology Internship Scheme. The quality improvement project was supported by the North Bristol NHS Trust and the "Next Big Thing Award," funded by the Southmead Hospital Charity. The views expressed in this publication are those of the authors and do not necessarily reflect those of the North Bristol NHS Trust, the UK NHS, or the NIHR. The authors would like to thank Omaid Ahmadyur for supporting the recruitment of the participants.

Authors' Contributions

AGKM and JB developed the idea for this study with input on the conceptualization from HLB, KA, and RM. CH and AGKM prepared the protocol, which was reviewed, discussed, and approved by all coauthors. CH, AGKM, KA, and RM established and formulated the methods for this review, with inputs from JB and VS. Recruitment and data collection were undertaken by CH, DH, SH, and AGKM. Data analysis was performed by CH and TD under the guidance of AGKM and RM. AGKM provided general oversight for this study. AGKM, SH, DH, and JB will take the lead on implementing the findings from this study with the help of CH, VS, and the ALPACA Study Team. The ALPACA Study Team members are: Andy Judge, Andrew Smith, Archana Lingampalli, Barnaby Reeves, Jessica Preshaw, Michael R Whitehouse, Paul Cresswell, Philip Braude, Shelley Potter, Timothy Beckitt, and Timothy Whittlestone.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Example screenshots of the measurement system.
[PDF File (Adobe PDF File), 1154 KB - humanfactors_v11i1e46698_app1.pdf]

Multimedia Appendix 2
Topic guide.
[PDF File (Adobe PDF File), 191 KB - humanfactors_v11i1e46698_app2.pdf]

Multimedia Appendix 3
Process map of tasks.
[PDF File (Adobe PDF File), 89 KB - humanfactors_v11i1e46698_app3.pdf]

References

1. The NHS constitution for England. Department of Health and Social Care. 2021. URL: https://www.gov.uk/government/publications/the-nhs-constitution-for-england/the-nhs-constitution-for-england#introduction-to-the-nhs-constitution [accessed 2022-09-15]
2. The NHS long term plan. National Health Service. 2019. URL: https://www.longtermplan.nhs.uk/ [accessed 2024-03-21]
3. Good surgical practice. Royal College of Surgeons. 2013. URL: https://www.rcseng.ac.uk/standards-and-research/gsp/ [accessed 2024-03-21]

XSL•FO
RenderX

4. Decision making and consent - ethical guidance. General Medical Council. 2020. URL: https://www.gmc-uk.org/professional-standards/professional-standards-for-doctors/decision-making-and-consent [accessed 2024-03-21]
5. Salzburg Global Seminar. Salzburg statement on shared decision making. BMJ 2011 Mar 22;342:d1745. [doi: 10.1136/bmj.d1745] [Medline: 21427038]
6. Elwyn G, Laitner S, Coulter A, Walker E, Watson P, Thomson R. Implementing shared decision making in the NHS. BMJ 2010 Oct 14;341:c5146 [FREE Full text] [doi: 10.1136/bmj.c5146] [Medline: 20947577]
7. Briss P, Rimer B, Reilley B, Coates RC, Lee NC, Mullen P, Task Force on Community Preventive Services. Promoting informed decisions about cancer screening in communities and healthcare systems. Am J Prev Med 2004 Jan;26(1):67-80. [doi: 10.1016/j.amepre.2003.09.012] [Medline: 14700715]
8. Goto Y, Miura H. Challenges in promoting shared decision-making: towards a breakthrough in Japan. Z Evid Fortbild Qual Gesundhwes 2022 Jun;171:84-88. [doi: 10.1016/j.zefq.2022.04.007] [Medline: 35610133]
9. Lu C, Li X, Yang K. Trends in shared decision-making studies from 2009 to 2018: a bibliometric analysis. Front Public Health 2019;7:384 [FREE Full text] [doi: 10.3389/fpubh.2019.00384] [Medline: 31921749]
10. Hughes TM, Merath K, Chen Q, Sun S, Palmer E, Idrees JJ, et al. Association of shared decision-making on patient-reported health outcomes and healthcare utilization. Am J Surg 2018 Jul;216(1):7-12. [doi: 10.1016/j.amjsurg.2018.01.011] [Medline: 29395026]
11. Tai-Seale M, Elwyn G, Wilson CJ, Stults C, Dillon EC, Li M, et al. Enhancing shared decision making through carefully designed interventions that target patient and provider behavior. Health Aff (Millwood) 2016 Apr;35(4):605-612. [doi: 10.1377/hlthaff.2015.1398] [Medline: 27044959]
12. Kashaf MS, McGill E. Does shared decision making in cancer treatment improve quality of life? A systematic literature review. Med Decis Making 2015 Nov;35(8):1037-1048. [doi: 10.1177/0272989X15598529] [Medline: 26246515]
13. Stacey D, Bennett CL, Barry MJ, Col NF, Eden KB, Holmes-Rovner M, et al. Decision aids for people facing health treatment or screening decisions. Cochrane Database Syst Rev 2011 Oct 05(10):CD001431. [doi: 10.1002/14651858.CD001431.pub3] [Medline: 21975733]
14. Durand M, Carpenter L, Dolan H, Bravo P, Mann M, Bunn F, et al. Do interventions designed to support shared decision-making reduce health inequalities? A systematic review and meta-analysis. PLoS One 2014;9(4):e94670 [FREE Full text] [doi: 10.1371/journal.pone.0094670] [Medline: 24736389]
15. Shay LA, Lafata JE. Where is the evidence? A systematic review of shared decision making and patient outcomes. Med Decis Making 2015 Jan;35(1):114-131 [FREE Full text] [doi: 10.1177/0272989X14551638] [Medline: 25351843]
16. Weiser TG, Haynes AB, Molina G, Lipsitz SR, Esquivel MM, Uribe-Leitz T, et al. Estimate of the global volume of surgery in 2012: an assessment supporting improved health outcomes. Lancet 2015 Apr 27;385 Suppl 2:S11. [doi: 10.1016/S0140-6736(15)60806-6] [Medline: 26313057]
17. Abbott TE, Fowler AJ, Dobbs TD, Harrison EM, Gillies MA, Pearse RM. Frequency of surgical treatment and related hospital procedures in the UK: a national ecological study using hospital episode statistics. Br J Anaesth 2017 Aug 01;119(2):249-257 [FREE Full text] [doi: 10.1093/bja/aex137] [Medline: 28854546]
18. Davison BJ, Matthew A, Gardner AM. Prospective comparison of the impact of robotic-assisted laparoscopic radical prostatectomy versus open radical prostatectomy on health-related quality of life and decision regret. Can Urol Assoc J 2014;8(1-2):E68-E72 [FREE Full text] [doi: 10.5489/cuaj.480] [Medline: 24454607]
19. Diefenbach MA, Mohamed NE. Regret of treatment decision and its association with disease-specific quality of life following prostate cancer treatment. Cancer Invest 2007 Sep 11;25(6):449-457. [doi: 10.1080/07357900701359460] [Medline: 17882657]
20. Walsh T, Barr PJ, Thompson R, Ozanne E, O'Neill C, Elwyn G. Undetermined impact of patient decision support interventions on healthcare costs and savings: systematic review. BMJ 2014 Jan 23;348:g188 [FREE Full text] [doi: 10.1136/bmj.g188] [Medline: 24458654]
21. Kissane DW, Bylund CL, Banerjee SC, Bialer PA, Levin TT, Maloney EK, et al. Communication skills training for oncology professionals. J Clin Oncol 2012 Apr 10;30(11):1242-1247 [FREE Full text] [doi: 10.1200/JCO.2011.39.6184] [Medline: 22412145]
22. Ross L, Ashford AD, Bleechington SJ, Dark T, Erwin DO. Applicability of a video intervention to increase informed decision making for prostate-specific antigen testing. J Natl Med Assoc 2010 Mar;102(3):228-236 [FREE Full text] [doi: 10.1016/s0027-9684(15)30529-0] [Medline: 20355352]
23. Smith SK, Trevena L, Simpson JM, Barratt A, Nutbeam D, McCaffery KJ. A decision aid to support informed choices about bowel cancer screening among adults with low education: randomised controlled trial. BMJ 2010 Oct 26;341:c5370 [FREE Full text] [doi: 10.1136/bmj.c5370] [Medline: 20978060]
24. Joseph-Williams N, Lloyd A, Edwards A, Stobbart L, Tomson D, Macphail S, et al. Implementing shared decision making in the NHS: lessons from the MAGIC programme. BMJ 2017 Apr 18;357:j1744 [FREE Full text] [doi: 10.1136/bmj.j1744] [Medline: 28420639]
25. Shard decision making NG197. National Institute for Health and Care Excellence. 2021. URL: https://www.nice.org.uk/guidance/ng197 [accessed 2024-03-21]

XSL•FO
RenderX

26. Légaré F, Adekpedjou R, Stacey D, Turcotte S, Kryworuchko J, Graham ID, et al. Interventions for increasing the use of shared decision making by healthcare professionals. Cochrane Database Syst Rev 2018 Jul 19;7(7):CD006732 [FREE Full text] [doi: 10.1002/14651858.CD006732.pub4] [Medline: 30025154]
27. Ubbink DT, Hageman MG, Legemate DA. Shared decision-making in surgery. Surg Technol Int 2015 May;26:31-36. [Medline: 26054988]
28. Elwyn G, Scholl I, Tietbohl C, Mann M, Edwards AG, Clay C, et al. "Many miles to go …": a systematic review of the implementation of patient decision support interventions into routine clinical practice. BMC Med Inform Decis Mak 2013;13 Suppl 2(Suppl 2):S14 [FREE Full text] [doi: 10.1186/1472-6947-13-S2-S14] [Medline: 24625083]
29. Forcino RC, Meinders MJ, Engel JA, O'Malley AJ, Elwyn G. Routine patient-reported experience measurement of shared decision-making in the USA: a qualitative study of the current state according to frontrunners. BMJ Open 2020 Jun 07;10(6):e037087 [FREE Full text] [doi: 10.1136/bmjopen-2020-037087] [Medline: 32513894]
30. Scholl I, Hahlweg P, Lindig A, Frerichs W, Zill J, Cords H, et al. Evaluation of a program for routine implementation of shared decision-making in cancer care: results of a stepped wedge cluster randomized trial. Implement Sci 2021 Dec 29;16(1):106 [FREE Full text] [doi: 10.1186/s13012-021-01174-4] [Medline: 34965881]
31. Elwyn G, Frosch DL, Kobrin S. Implementing shared decision-making: consider all the consequences. Implement Sci 2016 Aug 08;11:114 [FREE Full text] [doi: 10.1186/s13012-016-0480-9] [Medline: 27502770]
32. Härter M, van der Weijden T, Elwyn G. Policy and practice developments in the implementation of shared decision making: an international perspective. Z Evid Fortbild Qual Gesundhwes 2011;105(4):229-233. [doi: 10.1016/j.zefq.2011.04.018] [Medline: 21620313]
33. Zisman-Ilani Y, Chmielowska M, Dixon LB, Ramon S. NICE shared decision making guidelines and mental health: challenges for research, practice and implementation. BJPsych Open 2021 Sep 02;7(5):e154 [FREE Full text] [doi: 10.1192/bjo.2021.987] [Medline: 34470688]
34. Shared decision-making: summary guide. NHS England Personalised Care. URL: https://www.england.nhs.uk/publication/shared-decision-making-summary-guide/#:~:text=The%20Shared%20Decision%2DMaking%20Summary,services%20that%20informed%20individuals%20want [accessed 2024-03-21]
35. Tyack Z, Simons M, McPhail SM, Harvey G, Zappala T, Ware RS, et al. Improving the patient-centred care of children with life-altering skin conditions using feedback from electronic patient-reported outcome measures: protocol for a hybrid effectiveness-implementation study (PEDS-ePROM). BMJ Open 2021 Apr 09;11(4):e041861 [FREE Full text] [doi: 10.1136/bmjopen-2020-041861] [Medline: 33837095]
36. Kyte D, Anderson N, Auti R, Aiyegbusi OL, Bishop J, Bissell A, et al. Development of an electronic patient-reported outcome measure (ePROM) system to aid the management of patients with advanced chronic kidney disease. J Patient Rep Outcomes 2020 Jul 08;4(1):55 [FREE Full text] [doi: 10.1186/s41687-020-00223-8] [Medline: 32642867]
37. van der Krieke L, Emerencia AC, Aiello M, Sytema S. Usability evaluation of a web-based support system for people with a schizophrenia diagnosis. J Med Internet Res 2012 Feb 06;14(1):e24 [FREE Full text] [doi: 10.2196/jmir.1921] [Medline: 22311883]
38. Richards HS, Blazeby JM, Portal A, Harding R, Reed T, Lander T, et al. A real-time electronic symptom monitoring system for patients after discharge following surgery: a pilot study in cancer-related surgery. BMC Cancer 2020 Jun 10;20(1):543 [FREE Full text] [doi: 10.1186/s12885-020-07027-5] [Medline: 32522163]
39. Basch E, Deal AM, Kris MG, Scher HI, Hudis CA, Sabbatini P, et al. Symptom monitoring with patient-reported outcomes during routine cancer treatment: a randomized controlled trial. J Clin Oncol 2016 Feb 20;34(6):557-565 [FREE Full text] [doi: 10.1200/JCO.2015.63.0830] [Medline: 26644527]
40. Aiyegbusi OL, Kyte D, Cockwell P, Marshall T, Dutton M, Walmsley-Allen N, et al. Patient and clinician perspectives on electronic patient-reported outcome measures in the management of advanced CKD: a qualitative study. Am J Kidney Dis 2019 Aug;74(2):167-178. [doi: 10.1053/j.ajkd.2019.02.011] [Medline: 31003865]
41. van Gemert-Pijnen JE, Nijland N, van Limburg M, Ossebaard HC, Kelders SM, Eysenbach G, et al. A holistic framework to improve the uptake and impact of eHealth technologies. J Med Internet Res 2011 Dec 05;13(4):e111 [FREE Full text] [doi: 10.2196/jmir.1672] [Medline: 22155738]
42. O'Cathain A, Croot L, Sworn K, Duncan E, Rousseau N, Turner K, et al. Taxonomy of approaches to developing interventions to improve health: a systematic methods overview. Pilot Feasibility Stud 2019 Mar 12;5(1):41 [FREE Full text] [doi: 10.1186/s40814-019-0425-6] [Medline: 30923626]
43. Yardley L, Morrison L, Bradbury K, Muller I. The person-based approach to intervention development: application to digital health-related behavior change interventions. J Med Internet Res 2015 Jan 30;17(1):e30 [FREE Full text] [doi: 10.2196/jmir.4055] [Medline: 25639757]
44. Bevan N. Measuring usability as quality of use. Software Qual J 1995 Jun;4(2):115-130. [doi: 10.1007/BF00402715]
45. Maramba I, Chatterjee A, Newman C. Methods of usability testing in the development of eHealth applications: a scoping review. Int J Med Inform 2019 Jun;126:95-104. [doi: 10.1016/j.ijmedinf.2019.03.018] [Medline: 31029270]
46. Dumas JS, Salzman MC. Usability assessment methods. Rev Hum Factors Ergon 2016 Nov 03;2(1):109-140. [doi: 10.1177/1557234x0600200105]

XSL•FO
RenderX

47. Lin HX, Choong Y, Salvendy G. A proposed index of usability: a method for comparing the relative usability of different software systems. Behav Inf Technol 1997 Jan;16(4-5):267-277. [doi: 10.1080/014492997119833]
48. Yen PY, Bakken S. Review of health information technology usability study methodologies. J Am Med Inform Assoc 2012;19(3):413-422 [FREE Full text] [doi: 10.1136/amiajnl-2010-000020] [Medline: 21828224]
49. Aiyegbusi OL. Key methodological considerations for usability testing of electronic patient-reported outcome (ePRO) systems. Qual Life Res 2020 Feb;29(2):325-333 [FREE Full text] [doi: 10.1007/s11136-019-02329-z] [Medline: 31691202]
50. ISO 9241-11:2018: ergonomics of human-system interaction: part 11: usability: definitions and concepts. International Organization for Standardization. 2018. URL: https://www.iso.org/standard/63500.html#:~:text=ISO%209241%2D11%3A2018%20provides,services%20(including%20technical%20and%20personal [accessed 2024-03-21]
51. Bevan N, Carter J, Earthy J, Geis T, Harker S. New ISO standards for usability, usability reports and usability measures. In: Proceedings of the 18th International Conference on Human-Computer Interaction. Theory, Design, Development and Practice. 2016 Presented at: HCI '16; July 17-22, 2016; Toronto, ON p. 268-278 URL: https://link.springer.com/chapter/10.1007/978-3-319-39510-4_25 [doi: 10.1007/978-3-319-39510-4_25]
52. Hoffmann C, Avery KN, Macefield RC, Snelgrove V, Blazeby JM, Hopkins D, ALPACA Study team. Real-time monitoring and feedback to improve shared decision-making for surgery (the ALPACA Study): protocol for a mixed-methods study to inform co-development of an inclusive intervention. BMJ Open 2024 Jan 18;14(1):e079155 [FREE Full text] [doi: 10.1136/bmjopen-2023-079155] [Medline: 38238045]
53. Gärtner FR, Bomhof-Roordink H, Smith IP, Scholl I, Stiggelbout AM, Pieterse AH. The quality of instruments to assess the process of shared decision making: a systematic review. PLoS One 2018;13(2):e0191747 [FREE Full text] [doi: 10.1371/journal.pone.0191747] [Medline: 29447193]
54. Personalised care shared decision making summary guide. NHS England and NHS Improvement. 2019. URL: https://www.england.nhs.uk/wp-content/uploads/2019/01/shared-decision-making-summary-guide-v1.pdf [accessed 2024-03-21]
55. Barr PJ, Forcino RC, Thompson R, Ozanne EM, Arend R, Castaldo MG, et al. Evaluating CollaboRATE in a clinical setting: analysis of mode effects on scores, response rates and costs of data collection. BMJ Open 2017 Mar 24;7(3):e014681 [FREE Full text] [doi: 10.1136/bmjopen-2016-014681] [Medline: 28341691]
56. Doherr H, Christalle E, Kriston L, Härter M, Scholl I. Use of the 9-item shared decision making questionnaire (SDM-Q-9 and SDM-Q-Doc) in intervention studies-a systematic review. PLoS One 2017;12(3):e0173904 [FREE Full text] [doi: 10.1371/journal.pone.0173904] [Medline: 28358864]
57. Barr PJ, Thompson R, Walsh T, Grande SW, Ozanne EM, Elwyn G. The psychometric properties of CollaboRATE: a fast and frugal patient-reported measure of the shared decision-making process. J Med Internet Res 2014 Jan 03;16(1):e2 [FREE Full text] [doi: 10.2196/jmir.3085] [Medline: 24389354]
58. Barnum CM. Establishing the essentials. In: Barnum CM, editor. Usability Testing Essentials: Ready, Set ...Test!. Burlington, MA: Morgan Kaufmann; 2011:9-23.
59. Hedberg H, Iivari N, Rajanen M, Harjumaa L. Assuring quality and usability in open source software development. In: Proceedings of the 1st International Workshop on Emerging Trends in FLOSS Research and Development. 2007 Presented at: FLOSS '07; May 20-26, 2007; Minneapolis, MN URL: https://ieeexplore.ieee.org/document/4273073 [doi: 10.1109/floss.2007.2]
60. van den Haak MJ, de Jong MD, Schellens PJ. Evaluation of an informational web site: three variants of the think-aloud method compared. Tech Commun 2007;54(1):58-71 [FREE Full text]
61. Alshammari T, Alhadreti O, Mayhew PJ. When to ask participants to think aloud: a comparative study of concurrent and retrospective think-aloud methods. Int J Hum Comput Interact 2015;6(3):48-64 [FREE Full text] [doi: 10.4135/9781412961288.n460]
62. Khajouei R, Farahani F. A combination of two methods for evaluating the usability of a hospital information system. BMC Med Inform Decis Mak 2020 May 04;20(1):84 [FREE Full text] [doi: 10.1186/s12911-020-1083-6] [Medline: 32366248]
63. Farzanfar R, Finkelstein J, Friedman RH. Testing the usability of two automated home-based patient-management systems. J Med Syst 2004 Apr;28(2):143-153. [doi: 10.1023/b:joms.0000023297.50379.3c] [Medline: 15195845]
64. Kaufman DR, Patel VL, Hilliman C, Morin PC, Pevzner J, Weinstock RS, et al. Usability in the real world: assessing medical information technologies in patients' homes. J Biomed Inform 2003;36(1-2):45-60 [FREE Full text] [doi: 10.1016/s1532-0464(03)00056-x] [Medline: 14552846]
65. Birns J, Birns JH, Joffre KA, Leclerc JF, Paulsen CA. Getting the Whole Picture: Collecting Usability Data Using Two Methods-Concurrent Think Aloud and Retrospective Probing. 2002 Presented at: Proceedings of UPA Conference; 2-6 September 2002; London URL: https://citeseerx.ist.psu.edu/document?repid=rep1&type=pdf&doi=ba55a9ff388337e7e561fd1db41b3533051746a9 [doi: 10.1109/ipcc.2003.1245501]
66. Geisen E, Romano Bergstrom J. Developing the usability testing protocol. In: Geisen E, Romano Bergstrom J, editors. Usability Testing for Survey Research. Burlington, MA: Morgan Kaufmann; 2017:111-129.
67. Ghasemi A, Zahediasl S. Normality tests for statistical analysis: a guide for non-statisticians. Int J Endocrinol Metab 2012;10(2):486-489 [FREE Full text] [doi: 10.5812/ijem.3505] [Medline: 23843808]

XSL•FO
RenderX

68. Braun V, Clarke V. Using thematic analysis in psychology. Qual Res Psychol 2006 Jan;3(2):77-101. [doi: 10.1191/1478088706qp063oa]
69. Harris PA, Taylor R, Thielke R, Payne J, Gonzalez N, Conde JG. Research electronic data capture (REDCap)--a metadata-driven methodology and workflow process for providing translational research informatics support. J Biomed Inform 2009 Apr;42(2):377-381 [FREE Full text] [doi: 10.1016/j.jbi.2008.08.010] [Medline: 18929686]
70. Jaspers MW. A comparison of usability methods for testing interactive health technologies: methodological aspects and empirical evidence. Int J Med Inform 2009 May;78(5):340-353. [doi: 10.1016/j.ijmedinf.2008.10.002] [Medline: 19046928]
71. Virzi RA. Refining the test phase of usability evaluation: how many subjects is enough? Hum Factors 2016 Nov 23;34(4):457-468. [doi: 10.1177/001872089203400407]
72. Nielsen J, Landauer TK. Mathematical model of the finding of usability problems. In: Proceedings of the INTERACT '93 and CHI '93 Conference on Human Factors in Computing Systems. 1993 Presented at: CHI '93; April 24-29, 1993; Amsterdam, The Netherlands p. 206-213 URL: https://dl.acm.org/doi/10.1145/169059.169166 [doi: 10.1145/169059.169166]
73. Brooke J. SUS: a 'quick and dirty' usability scale. In: Jordan PW, Thomas B, McClelland IL, Weerdmeester B, editors. Usability Evaluation In Industry. London, UK: CRC Press; 1996.
74. Schnall R, Cho H, Liu J. Health information technology usability evaluation scale (Health-ITUES) for usability assessment of mobile health technology: validation study. JMIR Mhealth Uhealth 2018 Jan 05;6(1):e4 [FREE Full text] [doi: 10.2196/mhealth.8851] [Medline: 29305343]
75. Yen PY, Bakken S. Review of health information technology usability study methodologies. J Am Med Inform Assoc 2012;19(3):413-422 [FREE Full text] [doi: 10.1136/amiajnl-2010-000020] [Medline: 21828224]
76. O'Connell S, Palmer R, Withers K, Saha N, Puntoni S, Carolan-Rees G, PROMs, PREMsEffectiveness Programme. Requirements for the collection of electronic PROMS either "in clinic" or "at home" as part of the PROMs, PREMs and Effectiveness Programme (PPEP) in Wales: a feasibility study using a generic PROM tool. Pilot Feasibility Stud 2018;4:90 [FREE Full text] [doi: 10.1186/s40814-018-0282-8] [Medline: 29997899]
77. Palmen LN, Schrier JC, Scholten R, Jansen JH, Koëter S. Is it too early to move to full electronic PROM data collection?: A randomized controlled trial comparing PROM's after hallux valgus captured by e-mail, traditional mail and telephone. Foot Ankle Surg 2016 Mar;22(1):46-49. [doi: 10.1016/j.fas.2015.05.001] [Medline: 26869500]
78. Kyte D, Ives J, Draper H, Calvert M. Current practices in patient-reported outcome (PRO) data collection in clinical trials: a cross-sectional survey of UK trial staff and management. BMJ Open 2016 Oct 03;6(10):e012281 [FREE Full text] [doi: 10.1136/bmjopen-2016-012281] [Medline: 27697875]
79. Liu TC, Ohueri CW, Schryver EM, Bozic KJ, Koenig KM. Patient-identified barriers and facilitators to pre-visit patient-reported outcomes measures completion in patients with hip and knee pain. J Arthroplasty 2018 Mar;33(3):643-9.e1. [doi: 10.1016/j.arth.2017.10.022] [Medline: 29169689]
80. AOA PROMs pilot project final report. Australian Orthopaedic Association National Joint Replacement Registry (AOANJRR). 2020. URL: https://aoanjrr.sahmri.com/documents/10180/681914/AOANJRR+PROMs+Pilot+Final+Report [accessed 2024-03-21]
81. Lungu DA, Pennucci F, De Rosis S, Romano G, Melfi F. Implementing successful systematic patient reported outcome and experience measures (PROMs and PREMs) in robotic oncological surgery-the role of physicians. Int J Health Plann Manage 2020 May;35(3):773-787. [doi: 10.1002/hpm.2959] [Medline: 31793689]
82. Arner M. Developing a national quality registry for hand surgery: challenges and opportunities. EFORT Open Rev 2016 Apr;1(4):100-106 [FREE Full text] [doi: 10.1302/2058-5241.1.000045] [Medline: 28461935]
83. Morris ME, Brusco N, Woods J, Myles PS, Hodge A, Jones C, et al. Protocol for implementation of the 'AusPROM' recommendations for elective surgery patients: a mixed-methods cohort study. BMJ Open 2021 Sep 16;11(9):e049937 [FREE Full text] [doi: 10.1136/bmjopen-2021-049937] [Medline: 34531213]
84. Abbasi F, Khajouei R, Mirzaee M. The efficiency and effectiveness of surgery information systems in Iran. BMC Med Inform Decis Mak 2020 Sep 16;20(1):229 [FREE Full text] [doi: 10.1186/s12911-020-01236-5] [Medline: 32938452]
85. Szajna B. Empirical evaluation of the revised technology acceptance model. Manag Sci 1996 Jan;42(1):85-92. [doi: 10.1287/mnsc.42.1.85]
86. Kalankesh LR, Nasiry Z, Fein RA, Damanabi S. Factors influencing user satisfaction with information systems: a systematic review. Galen Med J 2020;9:e1686. [doi: 10.31661/gmj.v9i0.1686] [Medline: 34466567]
87. Aggelidis VP, Chatzoglou PD. Hospital information systems: measuring end user computing satisfaction (EUCS). J Biomed Inform 2012 Jun;45(3):566-579 [FREE Full text] [doi: 10.1016/j.jbi.2012.02.009] [Medline: 22426283]
88. Mustafa M, Alzubi S. Factors affecting the success of internet of things for enhancing quality and efficiency implementation in hospitals sector in Jordan during the crises of COVID-19. In: Chakraborty C, Banerjee A, Garg L, Rodrigues JJ, editors. Internet of Medical Things for Smart Healthcare: COVID-19 Pandemic. Singapore, Singapore: Springer; 2020:107-140.
89. Handayani PW, Hidayanto AN, Budi I. User acceptance factors of hospital information systems and related technologies: systematic review. Inform Health Soc Care 2018 Dec;43(4):401-426. [doi: 10.1080/17538157.2017.1353999] [Medline: 28829650]
90. Sidani S, Epstein DR, Bootzin RR, Moritz P, Miranda J. Assessment of preferences for treatment: validation of a measure. Res Nurs Health 2009 Aug;32(4):419-431 [FREE Full text] [doi: 10.1002/nur.20329] [Medline: 19434647]

XSL•FO
RenderX

91. Ahmed S, Zidarov D, Eilayyan O, Visca R. Prospective application of implementation science theories and frameworks to inform use of PROMs in routine clinical care within an integrated pain network. Qual Life Res 2021 Nov;30(11):3035-3047 [FREE Full text] [doi: 10.1007/s11136-020-02600-8] [Medline: 32876812]
92. Meirte J, Hellemans N, Anthonissen M, Denteneer L, Maertens K, Moortgat P, et al. Benefits and disadvantages of electronic patient-reported outcome measures: systematic review. JMIR Perioper Med 2020 Apr 03;3(1):e15588 [FREE Full text] [doi: 10.2196/15588] [Medline: 33393920]
93. Schamber EM, Takemoto SK, Chenok KE, Bozic KJ. Barriers to completion of patient reported outcome measures. J Arthroplasty 2013 Oct;28(9):1449-1453. [doi: 10.1016/j.arth.2013.06.025] [Medline: 23890831]
94. Jordan JE, Buchbinder R, Briggs AM, Elsworth GR, Busija L, Batterham R, et al. The health literacy management scale (HeLMS): a measure of an individual's capacity to seek, understand and use health information within the healthcare setting. Patient Educ Couns 2013 May;91(2):228-235. [doi: 10.1016/j.pec.2013.01.013] [Medline: 23419326]
95. Knight J, Ayyash K, Colling K, Dhesi J, Ewan V, Danjoux G, et al. A cohort study investigating the relationship between patient reported outcome measures and pre-operative frailty in patients with operable, non-palliative colorectal cancer. BMC Geriatr 2020 Aug 27;20(1):311 [FREE Full text] [doi: 10.1186/s12877-020-01715-4] [Medline: 32854632]
96. Mason JD, Blencowe NS, McNair AG, Stevens DJ, Avery KN, Pullyblank AM, et al. Investigating the collection and assessment of patient-reported outcome data amongst unplanned surgical hospital admissions: a feasibility study. Pilot Feasibility Stud 2015;1:16 [FREE Full text] [doi: 10.1186/s40814-015-0011-5] [Medline: 27965795]
97. Kwong E, Black N. Feasibility of collecting retrospective patient reported outcome measures (PROMs) in emergency hospital admissions. J Patient Rep Outcomes 2018 Nov 15;2(1):54 [FREE Full text] [doi: 10.1186/s41687-018-0077-y] [Medline: 30467820]
98. Iversen HH, Holmboe O, Bjertnaes O. Patient-reported experiences with general practitioners: a randomised study of mail and web-based approaches following a national survey. BMJ Open 2020 Oct 13;10(10):e036533 [FREE Full text] [doi: 10.1136/bmjopen-2019-036533] [Medline: 33051230]
99. Bliddal S, Banasik K, Pedersen OB, Nissen J, Cantwell L, Schwinn M, et al. Acute and persistent symptoms in non-hospitalized PCR-confirmed COVID-19 patients. Sci Rep 2021 Jun 23;11(1):13153 [FREE Full text] [doi: 10.1038/s41598-021-92045-x] [Medline: 34162913]
100. Wintner LM, Giesinger JM, Zabernigg A, Rumpold G, Sztankay M, Oberguggenberger AS, et al. Evaluation of electronic patient-reported outcome assessment with cancer patients in the hospital and at home. BMC Med Inform Decis Mak 2015 Dec 23;15:110 [FREE Full text] [doi: 10.1186/s12911-015-0230-y] [Medline: 26699708]
101. Aiyegbusi OL, Nair D, Peipert JD, Schick-Makaroff K, Mucsi I. A narrative review of current evidence supporting the implementation of electronic patient-reported outcome measures in the management of chronic diseases. Ther Adv Chronic Dis 2021;12:20406223211015958 [FREE Full text] [doi: 10.1177/20406223211015958] [Medline: 34104376]
102. Hutchings A, Neuburger J, Grosse Frie K, Black N, van der Meulen J. Factors associated with non-response in routine use of patient reported outcome measures after elective surgery in England. Health Qual Life Outcomes 2012 Mar 30;10:34 [FREE Full text] [doi: 10.1186/1477-7525-10-34] [Medline: 22462512]
103. Bevans M, Ross A, Cella D. Patient-reported outcomes measurement information system (PROMIS): efficient, standardized tools to measure self-reported health and quality of life. Nurs Outlook 2014;62(5):339-345 [FREE Full text] [doi: 10.1016/j.outlook.2014.05.009] [Medline: 25015409]
104. Wild D, Grove A, Martin M, Eremenco S, McElroy S, Verjee-Lorenz A, ISPOR Task Force for TranslationCultural Adaptation. Principles of good practice for the translation and cultural adaptation process for patient-reported outcomes (PRO) measures: report of the ISPOR task force for translation and cultural adaptation. Value Health 2005;8(2):94-104 [FREE Full text] [doi: 10.1111/j.1524-4733.2005.04054.x] [Medline: 15804318]
105. Rao JK, Anderson LA, Inui TS, Frankel RM. Communication interventions make a difference in conversations between physicians and patients: a systematic review of the evidence. Med Care 2007 Apr;45(4):340-349. [doi: 10.1097/01.mlr.0000254516.04961.d5] [Medline: 17496718]
106. Ivers N, Jamtvedt G, Flottorp S, Young JM, Odgaard-Jensen J, French SD, et al. Audit and feedback: effects on professional practice and healthcare outcomes. Cochrane Database Syst Rev 2012 Jun 13(6):CD000259. [doi: 10.1002/14651858.CD000259.pub3] [Medline: 22696318]
107. Pattinson RC, Say L, Makin JD, Bastos MH. Critical incident audit and feedback to improve perinatal and maternal mortality and morbidity. Cochrane Database Syst Rev 2005 Oct 19;2005(4):CD002961 [FREE Full text] [doi: 10.1002/14651858.CD002961.pub2] [Medline: 16235307]
108. de Jong RK, Snoek H, Staal WG, Klip H. The effect of patients' feedback on treatment outcome in a child and adolescent psychiatric sample: a randomized controlled trial. Eur Child Adolesc Psychiatry 2019 Jun;28(6):819-834 [FREE Full text] [doi: 10.1007/s00787-018-1247-4] [Medline: 30390148]
109. Skivington K, Matthews L, Simpson SA, Craig P, Baird J, Blazeby JM, et al. A new framework for developing and evaluating complex interventions: update of Medical Research Council guidance. BMJ 2021 Sep 30;374:n2061 [FREE Full text] [doi: 10.1136/bmj.n2061] [Medline: 34593508]

XSL•FO
RenderX

Abbreviations

ePRM: electronic patient-reported measure
ISO: International Organization for Standardization
ISRCTN: International Standard Randomised Controlled Trial Number
NHS: National Health Service
REDCap: Research Electronic Data Capture
SDM: shared decision-making

Edited by A Kushniruk; submitted 21.02.23; peer-reviewed by D Lungu, S Scott, S Azadnajafabad; comments to author 04.08.23; revised version received 02.10.23; accepted 02.03.24; published 10.04.24.

Please cite as:
Hoffmann C, Avery K, Macefield R, Dvořák T, Snelgrove V, Blazeby J, Hopkins D, Hickey S, Gibbison B, Rooshenas L, Williams A, Aning J, Bekker HL, McNair AGK, The ALPACA Study Team
Usability of an Automated System for Real-Time Monitoring of Shared Decision-Making for Surgery: Mixed Methods Evaluation
JMIR Hum Factors 2024;11:e46698
URL: https://humanfactors.jmir.org/2024/1/e46698
doi:10.2196/46698
PMID:38598276

©Christin Hoffmann, Kerry Avery, Rhiannon Macefield, Tadeáš Dvořák, Val Snelgrove, Jane Blazeby, Della Hopkins, Shireen Hickey, Ben Gibbison, Leila Rooshenas, Adam Williams, Jonathan Aning, Hilary L Bekker, Angus GK McNair, The ALPACA Study Team. Originally published in JMIR Human Factors (https://humanfactors.jmir.org), 10.04.2024. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on https://humanfactors.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Influence of Disease-Related Stigma on Patients' Decisions to Upload Medical Reports to the German Electronic Health Record: Randomized Controlled Trial

Niklas von Kalckreuth[1], MSc; Markus A Feufel[1], MSc, PhD

Division of Ergonomics, Department of Psychology and Ergonomics, Technische Universität Berlin, Berlin, Germany

Corresponding Author:
Niklas von Kalckreuth, MSc
Division of Ergonomics
Department of Psychology and Ergonomics
Technische Universität Berlin
Straße des 17 Juni 135
Berlin, 10623
Germany
Phone: 49 3031470747
Email: niklas.vkalckreuth@tu-berlin.de

Abstract

Background: The rollout of the electronic health record (EHR) represents a central component of the digital transformation of the German health care system. Although the EHR promises more effective, safer, and faster treatment of patients from a systems perspective, the successful implementation of the EHR largely depends on the patient. In a recent survey, 3 out of 4 Germans stated that they intend to use the EHR, whereas other studies show that the intention to use a technology is not a reliable and sufficient predictor of actual use.

Objective: Controlling for patients' intention to use the EHR, we investigated whether disease-specific risk perceptions related to the time course of the disease and disease-related stigma explain the additional variance in patients' decisions to upload medical reports to the EHR.

Methods: In an online user study, 241 German participants were asked to interact with a randomly assigned medical report that varied systematically in terms of disease-related stigma (high vs low) and disease time course (acute vs chronic) and to decide whether to upload it to the EHR.

Results: Disease-related stigma (odds ratio 0.154, $P<.001$) offset the generally positive relationship between intention to use and the upload decision (odds ratio 2.628, $P<.001$), whereas the disease time course showed no effect.

Conclusions: Even if patients generally intend to use the EHR, risk perceptions such as those related to diseases associated with social stigma may deter people from uploading related medical reports to the EHR. To ensure the reliable use of this key technology in a digitalized health care system, transparent and easy-to-comprehend information about the safety standards of the EHR are warranted across the board, even for populations that are generally in favor of using the EHR.

(JMIR Hum Factors 2024;11:e52625) doi:10.2196/52625

KEYWORDS

electronic health record; EHR; technology acceptance; upload behavior; health-related stigma; intention to use; intention-behavior gap; medical reports; stigma; Germany; patient decision; digital transformation; implementation; risk; decision; risk perception; social stigma; safety

Introduction

Background

The digital transformation of health care promises safety and efficiency gains by connecting all players in a health care system [1-3]. One key technology to connect health professionals, insurance providers, and patients is the electronic health record (EHR), which will be implemented nationwide and mandatory for all patients in Germany starting on January 1, 2025. In the EHR, patients' medical data (eg, findings, diagnoses, therapies, vaccinations, discharge reports, emergency data, and medication

XSL•FO
RenderX

plans [4,5]) can be digitally documented, exchanged, and viewed [4,6]. Better coordination of health data can ultimately save costs in the health care system [7-9].

In Germany, the Patient Data Protection Act [10] mandates that it is ultimately the patient who controls the type of data that are stored and can be viewed in the EHR. Although a recent survey found that 3 out of 4 Germans state that they intend to use the EHR [11], its success ultimately depends on whether and under what circumstances it is actually *used* to store and share health data. As described below based on the available literature, intention to use is *not* a sufficient and reliable predictor of EHR use. Therefore, in this study, we sought to investigate to what extent intention to use predicts actual use and what additional factors may need to be taken into account to more reliably predict EHR use.

Related Work

The technology acceptance model (TAM) and its extensions such as the unified theory of acceptance and use of technology (UTAUT) assume a positive relationship between intention to use a technology (technology acceptance) and actual use [12-14]. In fact, empirical studies on social networks and online banking show that the greater the intention to use, the more likely the technology will actually be used. However, the same studies also show a statistical discrepancy between intention and behavior, as evidenced by the different variance (R^2) accounted for by the two constructs [15-17]. Questionnaire studies on this so-called "intention-behavior gap" suggest that intention is not a reliable predictor of behavior and consequently that other influencing factors must exist [18,19]. For instance, in the context of social media and electronic commerce, users often have massive privacy concerns to disclose their data and their intentions to use are generally low. Nonetheless, users tend to disclose their data if the benefits they expect from using the applications are sufficiently high [20]; this phenomenon is called the "privacy paradox" and has been confirmed repeatedly [15,20,21]. However, questionnaire studies on digital health technologies show no such paradox and more nuanced patterns. For health technologies, privacy concerns thus far either had no influence [22-24] or have been shown to have a systematic negative impact on intentions and actual technology use [25,26]. In summary, based on the available research, it is unclear to what extent intention to use predicts the actual use of digital health technologies such as the EHR. Theories of technology acceptance infer a direct, positive influence, whereas the results of various questionnaire studies suggest that other factors must play a role given the intention-behavior gap. Although the influence of a few technology-related factors (eg, controllability of data) on the intention to use an EHR have been investigated, a thorough investigation of disease-related factors has not yet been performed.

Methodologically, usage behavior has mostly been investigated using self-report questions about the frequency of use [15,16,27-29], which is associated with several limitations. First, frequency of use is only meaningful if the system is already established and widely used. In the case of new systems such as the EHR in Germany, frequency of use cannot be surveyed. Second, the actual context of use can be difficult to simulate in questionnaire studies, making it difficult to distinguish between intention and behavior [30]. Since the models of technology acceptance described above (ie, TAM and UTAUT) have been evaluated using questionnaires, they may not provide reliable insights into usage behavior in the context of the EHR.

Therefore, to investigate usage behavior regarding the EHR in Germany, we selected a different approach for this study. In terms of uploading behavior, we first identified two possible use cases: (1) users who are living with different acute as well as chronic diseases ("patients with multimorbidity" use case), enabling a direct comparison between different medical findings in terms of risks and benefits of uploading to the EHR; and (2) users who are healthy or have little to no preexisting conditions before they develop a chronic or acute disease ("patients with first contact" use case). To investigate these use cases, we developed and used an interactive prototype of the EHR (ie, a click dummy) to investigate factors influencing the EHR users' decision to upload medical reports. Compared to questionnaire studies, this approach has the advantage that the interaction with the click dummy is closer to a real interaction with the EHR, thereby increasing the ecological validity of behavioral measures [30]. To investigate the first use case, we used a mixed methods design where the experimental intervention was based on an interview study with potential EHR users [31]. The interview study showed that the time course of a disease (chronic vs acute) and disease-related stigma influence people's decisions to upload a medical report to the EHR. The following experiment showed that respondents were more likely to upload a medical report of a chronic disease to the EHR than to upload a report of an acute condition. In contrast, respondents were less likely to upload a report of a disease with high stigma. When a disease with high stigma had a chronic time course, reports were still uploaded. We here report the results of the second use case in which participants interacted with one medical report only.

Methods

Ethical Considerations

This study was approved by the Ethics Committee of the Department of Psychology and Ergonomics (Institut für Psychologie und Arbeitswissenschaft) at Technische Universität (TU) Berlin (tracking number: AWB_KAL_1_230311). Participants volunteered to participate in the survey and informed consent was required. On the first page of the survey, participants were told about the investigator, the study purpose, what data were to be collected during the study, and where and for how long they would be stored. Participants were informed about the duration of the survey (approximately 8 minutes) as well as the compensation for participation. Participants were compensated with €1.60 (US $1.75) for their time and thus according to minimum wage. The participants also had the possibility to download a PDF of the participant information on the first page.

Participants' personal data and responses were kept entirely anonymous and password-protected in the department's data vault. An anonymized data set from the study was made available to other researchers for further analysis with open access. The documentation and availability of the research data

XSL•FO
RenderX

collected during the study were managed using the TU repository "DepositOnce," adhering to the regulations for ensuring good scientific practice at TU Berlin, the guidelines of the "DepositOnce" internal research data repository, and data protection regulations. Compliance with these repository guidelines ensures the indexing and findability of the research data by third parties.

Participants

The study was conducted from May 9 to June 10, 2023. Based on an a priori power analysis for a logistic regression with three predictors as well as a false-positive rate α of .05 and a power of $1-\beta=0.80$, we aimed for a sample size of 186 participants. Individuals 18 years and older residing in Germany were eligible to participate in the study. Another prerequisite was that participants had no previous personal experience (own illness) with the diseases mentioned in the medical reports, as affected people deal with disease-related stigma differently than people who are not affected by the disease [32]. Sampling was conducted through Prolific [33], a click worker platform characterized by high data quality [34]. A total of 275 individuals participated in the study. The mean participation time was 9 minutes, 28 seconds (SD 3 minutes, 47 seconds) and the median was 8 minutes, 36 seconds.

Design

We used a 2×2 between-subject study design with the independent variables stigma (high vs low) and time course of illness (chronic vs acute). Stigma was operationalized as the risk that the medical findings could negatively affect the private, professional, or social life of the affected person. For this purpose, the medical reports related to personal lifestyle, as reflected in tests for sexually transmitted diseases [31,32]. The time course is a classification of diseases in terms of their duration. These can be either acute (diseases of short duration that come on quickly) or chronic (diseases that develop slowly or last for a longer time). The dependent variable was the decision to upload the medical report (ie, whether participants were willing to upload the medical findings to the EHR). Furthermore, the intention to use the EHR was included as a covariate.

Materials

The stimuli used in the study were realistic but specially created for the purpose of the study. The medical reports were provided by various hospitals and a medical association. To make the reports appear as realistic as possible, they were edited on the official document heads of these institutions (see Multimedia Appendix 1). In selecting the diseases, both the related stigma and time course were systematically varied. To reflect different disease-related stigma, which covered different risks for professional and social life [35-38], diseases were divided according to their low and high stigmatization potential. To reflect different time courses, diseases were divided according to an acute and chronic time course. Furthermore, diseases were selected to occur regardless of age so that they would be perceived as realistic diseases by an age-diverse sample. Table 1 shows the diseases used as stimuli, categorized by level of stigma potential and time course.

Table 1. Diseases used in the stimuli, categorized by level of stigma potential and time course.

Stigma potential	Acute disease	Chronic disease
Low	Fractured wrist	Type 1 diabetes
High	STD[a] (gonorrhea)	Depression

[a]STD: sexually transmitted disease.

The interface software FIGMA was used to create the click dummy, which was modeled after the mobile EHR app of a German health insurance company (BARMER)—the eCare app—to support a realistic interaction with an EHR. Specifically, the click dummy allowed participants to upload findings, grant or revoke permissions to view medical reports, and create medication plans. Only the "Upload report" function was used in this study.

We used LimeSurvey (version 3.28.3+220315) to create and conduct a 5-page online survey (see Multimedia Appendix 2). The EHR click dummy and the medical reports were incorporated into the survey using iFrame. LimeSurvey software was used to ensure that all questions had to be answered to complete the study and receive the compensation. As in the previous study investigating the first use case [31], in this study, we tested the effect of the independent variables by querying the perceived privacy risk and perceived benefit of uploading findings to the EHR as manipulation checks. Based on the results of this previous study [31], we assumed that high stigma would result in a high perceived privacy risk and a chronic time course would result in a high perceived benefit of uploading the medical report. Perceived privacy risk, perceived benefit, and intention to use were measured using a 7-point Likert scale ranging from 1 ("strongly disagree") to 7 ("strongly agree"). The decision to upload the finding to the EHR was measured using a dichotomous item (yes/no).

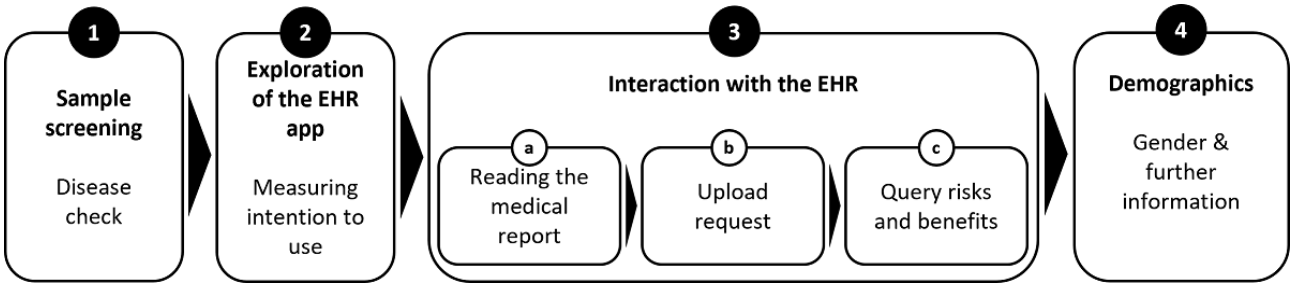
Procedure

The study procedure is shown schematically in Figure 1. Before the start of the experiment, participants gave their informed consent. This was followed by screening questions related to disease experience (step 1). Participants who had experience with the diseases in the medical reports were excluded from the study. Subsequently, participants were given 1 minute to interact with the EHR click dummy and were then required to answer questions regarding their intention to use the EHR (step 2). Participants were then asked to interact with the medical report (step 3). In this process, each person was first randomly assigned to one of the four diseases shown in Table 1 and asked to read an easy-to-understand description of the disease of approximately 2-3 sentences (see Multimedia Appendix 3) (step 3a). Participants then decided whether they wanted to upload

XSL•FO
RenderX

the report to their EHR (step 3b). Afterward, participants were asked to rate the perceived privacy risks and benefits of uploading the report (step 3c). The survey was completed with the collection of demographic characteristics (age, gender, education level, and experience with EHRs). In this step (step 4), the participants also had the opportunity to declare their responses invalid, while still receiving compensation, in case they did not pay sufficient attention to the instructions provided (eg, due to choosing random answers, inattentively reading questions, or rushing through the survey).

Figure 1. Overview of the study procedure. EHR: electronic health record.



Analysis

We cleaned and analyzed the data using RStudio (version 1.3.1093). Due to lack of variance inhomogeneity or a normal distribution, the analyses regarding perceived privacy risks and benefits were performed using the nonparametric Mann-Whitney *U* test. As mentioned above, we hypothesized that high stigma would result in a high perceived privacy risk and a chronic time course would result in a high perceived benefit of uploading the medical report. The influence of the independent variables (disease-specific stigma and time course) and the covariate "intention to use" on the upload decision were tested using multiple logistic regression with dummy coding. We hypothesized that usage behavior is negatively influenced by disease-specific stigma and positively influenced by time course and intention. To control for demographic and interindividual influences, we used multiple logistic regression with standardized coefficients for better comparability. In doing so, we followed the recommendations for testing control variables [39] and tested the variables that have been shown to be causally related to privacy behavior along with the independent variables. The control variables were age, education level, and experience with the technical system, in this case the EHR [40,41].

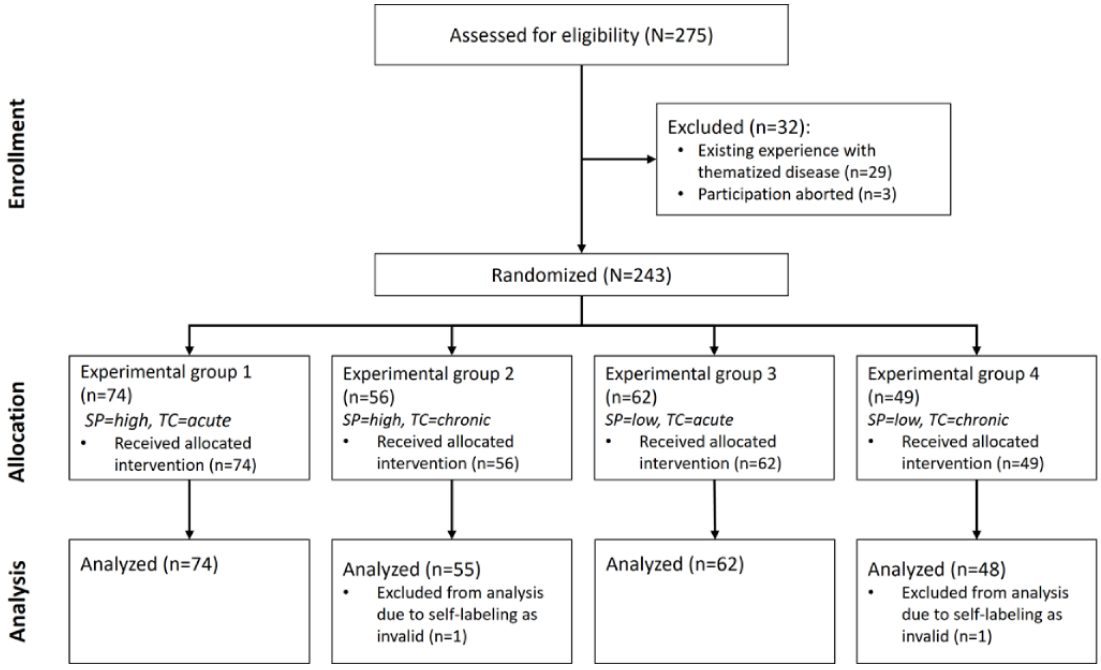
Results

Sample Characteristics

A total of 275 observations were collected. Of those, 34 records were excluded, 29 because of participants' previous medical histories, 3 because of incomplete questionnaires, and 2 because they were marked as invalid by participants. Figure 2 shows the flow of participants in the study based on the CONSORT (Consolidated Standards of Reporting Trials) statement [42].

Thus, a sample of 241 observations (146 male, 92 female, 1 diverse, 2 no information provided) was used for further analysis. Table 2 summarizes the demographic characteristics of the sample.

Figure 2. CONSORT (Consolidated Standards of Reporting Trials) flow chart. SP: stigma potential; TC: time course.



XSL•FO
RenderX

Table 2. Demographic data of the sample (N=241).

Demographic characteristic	Value
Age (years), mean (SD)	31.31 (9.76)
Gender, n (%)	
Female	92 (38.2)
Male	146 (60.6)
Other	3 (1.2)
Education, n (%)	
No degree	9 (3.7)
School leaving certificate	3 (1.2)
Secondary school certificate	18 (7.5)
General qualification for university entrance	66 (27.4)
Vocational training	33 (13.7)
University degree (bachelor's or master's degree)	112 (46.5)
Experience with the German EHR[a], n (%)	
EHR is unknown	61 (25.3)
EHR is known but not used	164 (68)
Occasional use	14 (5.8)
Regular use	2 (0.8)

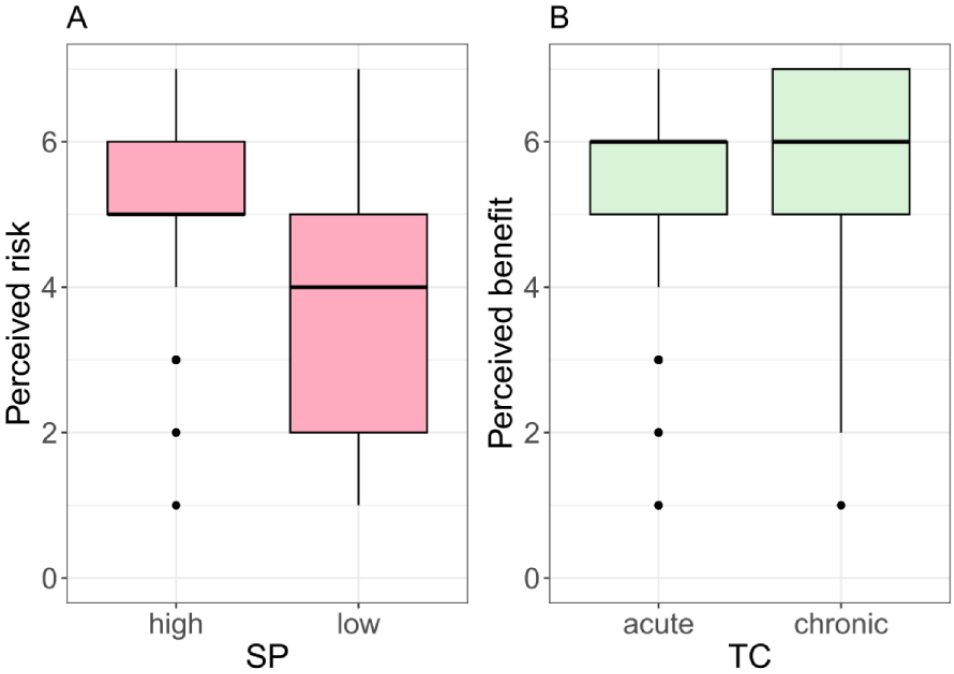
[a]EHR: electronic health record.

Risk and Benefit Perception

We first checked whether stigma potential had an effect on privacy risk perception and whether time course had an effect on the benefit perception of uploading (see Figure 3). Mann-Whitney *U* tests showed a significant effect of stigma potential on privacy risk perception (W=10,777; P<.001), where high stigma was associated with high risk. The effect of the disease time course on benefit perception was not significant (W=6379; P=.14), with a mean benefit perception of 5.34 (SD 1.39) for acute diseases and of 5.54 (SD 1.43) for chronic diseases. Consequently, in contrast to our study on the first use case with several medical reports [31], there was no relationship found between time course and perceived benefits when there is only one report to upload.

XSL•FO
RenderX

Figure 3. (A) Perceived risk in relation to stigma potential (SP) and (B) perceived benefit in relation to the disease time course (TC). The horizontal line in the box represents the median.



Controls

To investigate the potential association between the decision to upload the medical report and the independent variables disease-specific stigma and time course, we first performed a logistic regression (Hosmer-Lemeshow R^2=0.319, Nagelkerke R^2=0.590, Cox-Snell R^2=0.537; χ^2_{15}=86.973; P<.001) to control for the covariate intention to use and the demographic variables age, sex, education level, and experience with the EHR. The covariate intention to use (odds ratio [OR] 2.497, 95% CI 1.831-3.456; z=5.455; P<.001) showed an association with the decision to upload, whereas none of the control variables had an effect. These variables were consequently removed from the model for further analyses.

Uploading Behavior

To examine the association between the decision to upload and the independent variables stigma potential and time course, we performed a logistic regression controlling for the covariate intention to use (Hosmer-Lemeshow R^2=0.289, Nagelkerke R^2=0.551, Cox-Snell R^2=0.501; χ^2_3=78.748; P<.001). Intention to use was positively associated with uploading behavior; specifically, as intention to use increased, it was more than twice as likely that the report was uploaded to the EHR. In addition, there was a negative association between stigma and the decision to upload; specifically, when stigma was high, it was six times less likely that the report was uploaded than when stigma was low. Time course of the disease was not associated with the decision to upload a report. The summary of the results of the logistic regression are shown in Table 3.

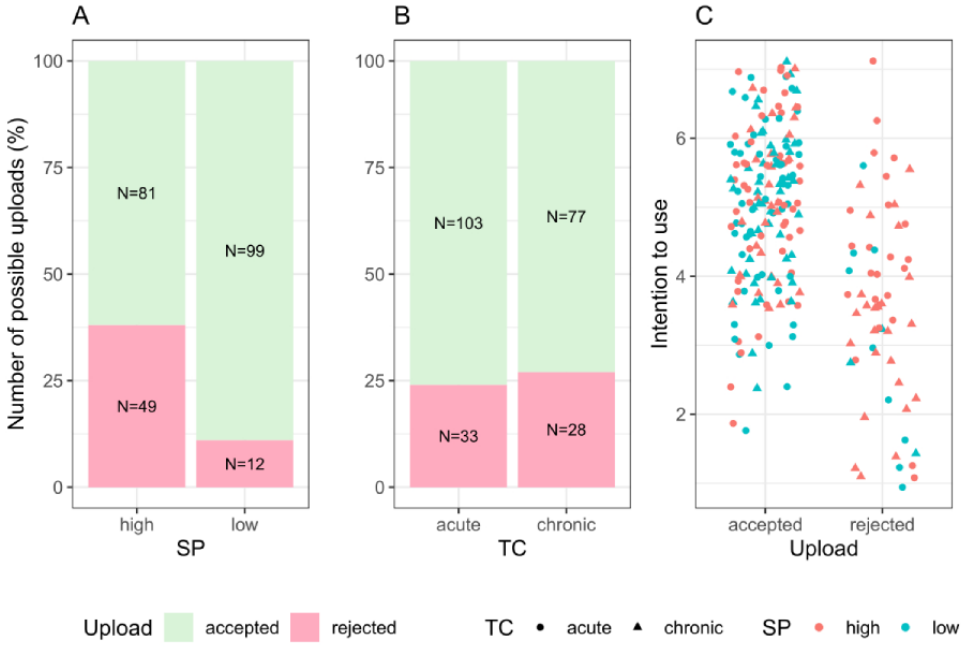
Table 3. Results of the logistic regression.

Variable	z value	P value	Odds ratio (95% CI)
Intention to use	6.210	<.001	2.682 (1.971-3.639)
Stigma potential	4.463	<.001	0.154 (0.064-0.336)
Time course	0.244	.81	1.093 (0.537-2.254)

The number of uploads is shown in Figure 4 in relation to the independent variables stigma potential (Figure 4A) and time course (Figure 4B). In addition, we show the relationship between intention to use and the decision to upload a report as a function of the independent variables in Figure 4C.

XSL•FO
RenderX

Figure 4. Number of uploads to the electronic health record with respect to (A) stigma potential (SP) and (B) time course (TC), and (C) the influence of intention to use on the decision to upload as a function of SP and TC.



Discussion

Principal Findings

Our results show that the decision to upload an individual medical report is influenced by people's intention to use the EHR. However, the stigma potential of the diseases mentioned in the reports also influenced this decision. Specifically, uploading diseases with high stigma was associated with higher privacy risk than diseases with low stigma (see Figure 3A). Consequently, stigma potential had a negative influence on the decision to upload records (see Figure 4A), despite generally high intentions to use the EHR.

Thus, intention to use predicts the use of the EHR in part, whereas disease-specific factors such as related stigma can override the general intention. This is particularly evident in Figure 4C where the participants who uploaded reports both with high and low stigma had mostly high intentions to use the EHR (scores>4). However, such a clear distribution of intention to use (scores<4) did not emerge in the case of rejection of uploading. Rather, it is notable that the rejected findings are mainly those with high stigmatization potential (majority of red dots/triangles in Figure 4C). This shows that the effect strength of the stigmatization potential (OR 0.154) is significantly greater than that of the intention to use (OR 2.628). The fact that uploading is rejected due to disease-specific stigma despite high intention to use supports the assumptions of an intention-behavior gap in EHR use [18].

The time course of the diseases had no influence on the decision to upload an individual record. Findings with chronic and acute diseases were uploaded by the majority of participants and with approximately equal frequency (see Figure 3B).

For both use cases, case 1 (patients with multimorbidity) and case 2 (patients after first contact), the results suggest that disease-specific stigma seems to exert an inhibiting influence on the decision to upload. In contrast, the time course only played a role in use case 1, where people interact with multiple reports at a time [31], but not when they interact with only one medical report (use case 2). This difference may be explained by the fact that patients' "health concerns" have a positive influence on their intention to share health data with others [22]. When faced with multiple medical reports, patients may be more aware and concerned about interactions between chronic diseases, because they more strongly affect the patient's health both now and in the future; consequently, the willingness to upload reports about chronic diseases increases. With a single report, interactions between diseases are less present, which means that the time course of a disease may play a reduced role in the decision to upload a record.

Implications

Both the intention to use and the stigma potential of diseases seem to influence whether patients upload an individual medical report to the EHR. Thus, in addition to increasing people's general intention to use the EHR via marketing and information, transparent and easy-to-comprehend information about the safety standards of the EHR (eg, for encrypting data) and the protection of medical records (eg, the control of access rights) are warranted, even for populations that are already in favor of using the EHR. Such combined interventions may help to reduce security concerns and enable realistic risk assessments of a data leak to ultimately ensure reliable use of the EHR as a key technology in any digitalized health care system.

Limitations and Future Directions

We deliberately excluded participants who already had a medical history with the diseases addressed in the stimuli to avoid bias in their responses. Individuals living with a stigmatized disease are more cautious to disclose the information, especially if the

XSL•FO
RenderX

disease is not immediately apparent [32,43]. The question arises to what extent the behavior of stigmatized individuals can be simulated under experimental conditions, provided that participants do not exhibit stigmatized characteristics. To further strengthen the validity and generalizability of our results, a follow-up study should examine the perspective of already affected individuals and compare the findings with the results of this study.

Another limitation is that the chronic and acute disease patterns used in the stimuli are not readily comparable. We decided to use the diseases listed in Table 1 as stimuli because they achieved the expected effects in the preliminary study [31]. We could only partially replicate these findings in the present study. For future studies, it would make sense to use diseases that can be more readily compared in terms of their stigma potential and time course (eg, gonorrhea and HIV or a wrist fracture and arthritis) to further strengthen the generalizability of the present findings.

Another limitation is that the distribution of our sample in terms of gender, age, and level of education does not correspond to that of the average German population. In particular, the level of education of our sample was above average. Although we were unable to detect any effects of the control variables age, gender, and level of education in the analysis, the results of this study should be validated with a more representative sample in the future.

Conclusions

In our study, we investigated which disease-specific factors influence whether medical reports are uploaded to the EHR in a German setting. To answer this question, we varied the stigma potential and the time course of diseases in medical reports and controlled for the influence of participants' intention to use the EHR on uploading behavior. We demonstrated that intention to use had a positive effect on the decision to upload a report. In addition, we found that the stigma potential of the disease listed on the medical reports can inhibit uploading behavior. In particular, we found that the intention to use the EHR may be offset by the stigma potential of a specific record.

In summary, despite the fact that 3 out of 4 Germans state that they intend to use the EHR [11], actual use of this technology may depend on disease-specific factors. Consequently, to ensure successful implementation of the EHR, stakeholders in the health system should not only promote the EHR per se but further develop formats and evaluate them with the help of user testing that provide transparent and easy-to-comprehend information about the standards of data security and control in the EHR. Only in this way can users realistically assess the risks associated with individual EHR use and make an informed decision for (or against) EHR use.

Acknowledgments

We acknowledge support from the German Research Foundation and the Open Access Publication Fund of Technische Universität Berlin. We also thank the eV Studienwerk Villigst and the German Federal Ministry of Education and Research, who provided the doctoral scholarship (to NvK) without which this research would not have been possible. We would also like to thank the institutions that provided us with medical findings as well as all those who participated in the study.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Medical findings used as stimuli.
[PDF File (Adobe PDF File), 377 KB - humanfactors_v11i1e52625_app1.pdf]

Multimedia Appendix 2
Questionnaire.
[DOCX File , 17 KB - humanfactors_v11i1e52625_app2.docx]

Multimedia Appendix 3
Disease descriptions.
[DOCX File , 13 KB - humanfactors_v11i1e52625_app3.docx]

Multimedia Appendix 4
CONSORT-EHEALTH checklist (V 1.6.1).
[PDF File (Adobe PDF File), 26097 KB - humanfactors_v11i1e52625_app4.pdf]

References

1. Haleem A, Javaid M, Pratap Singh R, Suman R. Medical 4.0 technologies for healthcare: features, capabilities, and applications. Internet of Things Cyber-Phys Syst 2022;2:12-30. [doi: 10.1016/j.iotcps.2022.04.001]

XSL•FO
RenderX

2. Hubmann M, Pätzmann-Sietas B, Morbach H. Telemedizin und digitale Akte – Wo stehen wir? Monatsschr Kinderheilkd 2021 Jul 09;169(8):711-716. [doi: 10.1007/s00112-021-01241-6]
3. Kraus S, Schiavone F, Pluzhnikova A, Invernizzi AC. Digital transformation in healthcare: analyzing the current state-of-research. J Bus Res 2021 Feb;123:557-567. [doi: 10.1016/j.jbusres.2020.10.030]
4. Bertram N, Püschner F, Gonçalves A, Binder S, Amelung V. Introduction of an electronic health record in Germany against the background of international experience. In: Klauber J, Geraedts M, Friedrich J, Wasem J, editors. Krankenhaus-Report 2019. Das digitale Krankenhaus. Berlin, Heidelberg: Springer; 2019:3-16.
5. Die elektronische Patientenakte (ePA). Bundesministerium für Gesundheit (BMG). 2021. URL: https://www.bundesgesundheitsministerium.de/elektronische-patientenakte.html [accessed 2023-04-14]
6. Elektronische Patientenakte - ePA. Kassenärztliche Bundesvereinigung (KBV). 2023. URL: https://www.kbv.de/html/epa.php [accessed 2023-04-14]
7. Baron von Osthoff M, Watzlaw-Schmidt U, Lehmann T, Hübner J. Patient group-specific data sovereignty needs and acceptance of electronic health records. Bundesgesundheitsblatt Gesundheitsforschung Gesundheitsschutz 2022 Nov 23;65(11):1197-1203 [FREE Full text] [doi: 10.1007/s00103-022-03589-w] [Medline: 36149472]
8. Cherif E, Mzoughi M. Electronic health record adopters: a typology based on patients' privacy concerns and perceived benefits. Public Health 2022 Jun;207:46-53. [doi: 10.1016/j.puhe.2022.03.010] [Medline: 35486983]
9. Pagliari C, Detmer D, Singleton P. Potential of electronic personal health records. BMJ 2007 Aug 18;335(7615):330-333 [FREE Full text] [doi: 10.1136/bmj.39279.482963.AD] [Medline: 17703042]
10. Patientendaten-Schutz-Gesetz. Bundesministerium für Gesundheit (BMG). 2020. URL: https://www.bundesgesundheitsministerium.de/patientendaten-schutz-gesetz.html [accessed 2023-04-14]
11. Elektronische patientenakte: geplante Wwderspruchslösung trifft auf breite zustimmung. Bertelsmann Stiftung. 2023. URL: https://tinyurl.com/mptd2j95 [accessed 2023-05-17]
12. Bettiga D, Lamberti L, Lettieri E. Individuals' adoption of smart technologies for preventive health care: a structural equation modeling approach. Health Care Manag Sci 2020 Jun 26;23(2):203-214. [doi: 10.1007/s10729-019-09468-2] [Medline: 30684067]
13. Davis FD. Perceived usefulness, perceived ease of use, and user acceptance of information technology. MIS Quart 1989 Sep;13(3):319-340. [doi: 10.2307/249008]
14. Venkatesh V, Morris MG, Davis GB, Davis FD. User acceptance of information technology: toward a unified view. MIS Quart 2003;27(3):425-478. [doi: 10.2307/30036540]
15. Kim B, Kim D. Understanding the key antecedents of users' disclosing behaviors on social networking sites: the privacy paradox. Sustainability 2020 Jun 24;12(12):5163. [doi: 10.3390/su12125163]
16. Kim B, Min J. The distinct roles of dedication-based and constraint-based mechanisms in social networking sites. Internet Res 2013;25(1):30-51. [doi: 10.1108/intr-11-2013-0253]
17. Oliveira T, Faria M, Thomas MA, Popovič A. Extending the understanding of mobile banking adoption: when UTAUT meets TTF and ITM. Int J Inf Manag 2014 Oct;34(5):689-703. [doi: 10.1016/j.ijinfomgt.2014.06.004]
18. Conner M, Norman P. Understanding the intention-behavior gap: the role of intention strength. Front Psychol 2022 Aug 4;13:923464 [FREE Full text] [doi: 10.3389/fpsyg.2022.923464] [Medline: 35992469]
19. Sheeran P, Webb TL. The intention–behavior gap. Soc Person Psych 2016 Sep 05;10(9):503-518. [doi: 10.1111/spc3.12265]
20. Kokolakis S. Privacy attitudes and privacy behaviour: a review of current research on the privacy paradox phenomenon. Comput Secur 2017 Jan;64:122-134. [doi: 10.1016/j.cose.2015.07.002]
21. Norberg PA, Horne DR, Horne DA. The privacy paradox: personal information disclosure intentions versus behaviors. J Consum Aff 2007 Mar 06;41(1):100-126. [doi: 10.1111/j.1745-6606.2006.00070.x]
22. Cherif E, Bezaz N, Mzoughi M. Do personal health concerns and trust in healthcare providers mitigate privacy concerns? Effects on patients' intention to share personal health data on electronic health records. Soc Sci Med 2021 Aug;283:114146. [doi: 10.1016/j.socscimed.2021.114146] [Medline: 34242891]
23. von Kalckreuth N, Feufel MA. Extending the privacy calculus to the mHealth domain: survey study on the intention to use mHealth apps in Germany. JMIR Hum Factors 2023 Aug 16;10:e45503 [FREE Full text] [doi: 10.2196/45503] [Medline: 37585259]
24. Zhu M, Wu C, Huang S, Zheng K, Young SD, Yan X, et al. Privacy paradox in mHealth applications: an integrated elaboration likelihood model incorporating privacy calculus and privacy fatigue. Telematics Informatics 2021 Aug;61:101601. [doi: 10.1016/j.tele.2021.101601]
25. Dinev T, Albano V, Xu H, D'Atri A, Hart P. Individuals' attitudes towards electronic health records: a privacy calculus perspective. In: Gupta A, Patel VL, Greenes RA, editors. Advances in Healthcare Informatics and Analytics. Cham: Springer International Publishing; 2016:19-50.
26. Li T, Slee T. The effects of information privacy concerns on digitizing personal health records. J Assn Inf Sci Tech 2014 Apr 21;65(8):1541-1554. [doi: 10.1002/asi.23068]
27. Kim SS, Malhotra NK, Narasimhan S. Research Note—Two competing perspectives on automatic use: a theoretical and empirical comparison. Inf Syst Res 2005 Dec;16(4):418-432. [doi: 10.1287/isre.1050.0070]

XSL•FO
RenderX

28. Pavlou PA. Consumer acceptance of electronic commerce: integrating trust and risk with the technology acceptance model. Int J Electron Commer 2014 Dec 23;7(3):101-134. [doi: 10.1080/10864415.2003.11044275]
29. Schretzlmaier P, Hecker A, Ammenwerth E. Extension of the Unified Theory of Acceptance and Use of Technology 2 model for predicting mHealth acceptance using diabetes as an example: a cross-sectional validation study. BMJ Health Care Inform 2022 Nov 15;29(1):e100640 [FREE Full text] [doi: 10.1136/bmjhci-2022-100640] [Medline: 36379608]
30. Distler V, Fassl M, Habib H, Krombholz K, Lenzini G, Lallemand C, et al. Empirical research methods in usable privacy and security. In: Gerber N, Stöver A, Marky K, editors. Human Factors in Privacy Research. Cham: Springer International Publishing; 2023:29-53.
31. von Kalckreuth N, Prümper AM, Feufel MA. The Influence of health data on the use of the electronic health record (EHR) -- a mixed methods approach. 2023 Presented at: AMCIS 2023; August 10-12, 2023; Panama City, Panama URL: https://aisel.aisnet.org/amcis2023/sig_health/sig_health/2
32. Goffman E. Stigma: Notes on the Management of Spoiled Identity. Saddle River, NJ: Prentice Hall; 1963.
33. Prolific. URL: https://www.prolific.co/ [accessed 2023-09-06]
34. Peer E, Brandimarte L, Samat S, Acquisti A. Beyond the Turk: alternative platforms for crowdsourcing behavioral research. J Exp Soc Psychol 2017 May;70:153-163. [doi: 10.1016/j.jesp.2017.01.006]
35. Glozier N. Workplace effects of the stigmatization of depression. J Occup Environ Med 1998 Sep;40(9):793-800. [doi: 10.1097/00043764-199809000-00008] [Medline: 9777563]
36. Strømsvik N, Råheim M, Oyen N, Engebretsen LF, Gjengedal E. Stigmatization and male identity: Norwegian males' experience after identification as BRCA1/2 mutation carriers. J Genet Couns 2010 Aug 20;19(4):360-370. [doi: 10.1007/s10897-010-9293-1] [Medline: 20306122]
37. Bickford J, Barton SE, Mandalia S. Chronic genital herpes and disclosure.... The influence of stigma. Int J STD AIDS 2007 Sep 25;18(9):589-592. [doi: 10.1258/095646207781568484] [Medline: 17784999]
38. Zacks S, Beavers K, Theodore D, Dougherty K, Batey B, Shumaker J, et al. Social stigmatization and hepatitis C virus infection. J Clin Gastroenterol 2006 Mar;40(3):220-224. [doi: 10.1097/00004836-200603000-00009] [Medline: 16633123]
39. Wysocki AC, Lawson KM, Rhemtulla M. Statistical control requires causal justification. Adv Methods Pract Psychol Sci 2022 Jun 28;5(2):251524592210958. [doi: 10.1177/25152459221095823]
40. Schomakers E, Lidynia C, Vervier LS, Calero Valdez A, Ziefle M. Applying an extended UTAUT2 model to explain user acceptance of lifestyle and therapy mobile health apps: survey study. JMIR Mhealth Uhealth 2022 Jan 18;10(1):e27095 [FREE Full text] [doi: 10.2196/27095] [Medline: 35040801]
41. Uncovska M, Freitag B, Meister S, Fehring L. Patient acceptance of prescribed and fully reimbursed mHealth apps in Germany: an UTAUT2-based online survey study. J Med Syst 2023 Jan 27;47(1):14 [FREE Full text] [doi: 10.1007/s10916-023-01910-x] [Medline: 36705853]
42. Cuschieri S. The CONSORT statement. Saudi J Anaesth 2019 Apr;13(Suppl 1):S27-S30 [FREE Full text] [doi: 10.4103/sja.SJA_559_18] [Medline: 30930716]
43. Weiss MG, Ramakrishna J, Somma D. Health-related stigma: rethinking concepts and interventions. Psychol Health Med 2006 Aug;11(3):277-287. [doi: 10.1080/13548500600595053] [Medline: 17130065]

Abbreviations

CONSORT: Consolidated Standards of Reporting Trails
EHR: electronic health record
OR: odds ratio
TAM: technology acceptance model
TU: Technische Universität
UTAUT: unified theory of acceptance and use of technology

Edited by P Santana-Mancilla; submitted 10.09.23; peer-reviewed by JJ Beunza, EM Schomakers; comments to author 07.02.24; revised version received 12.02.24; accepted 21.02.24; published 10.04.24.

Please cite as:
von Kalckreuth N, Feufel MA
Influence of Disease-Related Stigma on Patients' Decisions to Upload Medical Reports to the German Electronic Health Record: Randomized Controlled Trial
JMIR Hum Factors 2024;11:e52625
URL: https://humanfactors.jmir.org/2024/1/e52625
doi:10.2196/52625
PMID:38598271

XSL•FO
RenderX

©Niklas von Kalckreuth, Markus A Feufel. Originally published in JMIR Human Factors (https://humanfactors.jmir.org), 10.04.2024. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on https://humanfactors.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Academic Detailing as a Health Information Technology Implementation Method: Supporting the Design and Implementation of an Emergency Department–Based Clinical Decision Support Tool to Prevent Future Falls

Hanna J Barton[1], PhD; Apoorva Maru[1], BA, BSc; Margaret A Leaf[2], MSc; Daniel J Hekman[1], MSc; Douglas A Wiegmann[3], PhD; Manish N Shah[1], MD, MPH; Brian W Patterson[1,3], MD, MPH

[1]BerbeeWalsh Department of Emergency Medicine, University of Wisconsin-Madison, Madison, WI, United States

[2]Department of Information Services, UW Health, Madison, WI, United States

[3]Department of Industrial and Systems Engineering, University of Wisconsin-Madison, Madison, WI, United States

Corresponding Author:
Hanna J Barton, PhD
BerbeeWalsh Department of Emergency Medicine
University of Wisconsin-Madison
800 University Bay Dr.
Madison, WI, 53705
United States
Phone: 1 (608) 890 8682
Email: hbarton@wisc.edu

Abstract

Background: Clinical decision support (CDS) tools that incorporate machine learning–derived content have the potential to transform clinical care by augmenting clinicians' expertise. To realize this potential, such tools must be designed to fit the dynamic work systems of the clinicians who use them. We propose the use of academic detailing—personal visits to clinicians by an expert in a specific health IT tool—as a method for both ensuring the correct understanding of that tool and its evidence base and identifying factors influencing the tool's implementation.

Objective: This study aimed to assess academic detailing as a method for simultaneously ensuring the correct understanding of an emergency department–based CDS tool to prevent future falls and identifying factors impacting clinicians' use of the tool through an analysis of the resultant qualitative data.

Methods: Previously, our team designed a CDS tool to identify patients aged 65 years and older who are at the highest risk of future falls and prompt an interruptive alert to clinicians, suggesting the patient be referred to a mobility and falls clinic for an evidence-based preventative intervention. We conducted 10-minute academic detailing interviews (n=16) with resident emergency medicine physicians and advanced practice providers who had encountered our CDS tool in practice. We conducted an inductive, team-based content analysis to identify factors that influenced clinicians' use of the CDS tool.

Results: The following categories of factors that impacted clinicians' use of the CDS were identified: (1) aspects of the CDS tool's design (2) clinicians' understanding (or misunderstanding) of the CDS or referral process, (3) the busy nature of the emergency department environment, (4) clinicians' perceptions of the patient and their associated fall risk, and (5) the opacity of the referral process. Additionally, clinician education was done to address any misconceptions about the CDS tool or referral process, for example, demonstrating how simple it is to place a referral via the CDS and clarifying which clinic the referral goes to.

Conclusions: Our study demonstrates the use of academic detailing for supporting the implementation of health information technologies, allowing us to identify factors that impacted clinicians' use of the CDS while concurrently educating clinicians to ensure the correct understanding of the CDS tool and intervention. Thus, academic detailing can inform both real-time adjustments of a tool's implementation, for example, refinement of the language used to introduce the tool, and larger scale redesign of the CDS tool to better fit the dynamic work environment of clinicians.

(JMIR Hum Factors 2024;11:e52592) doi:10.2196/52592

XSL•FO
RenderX

KEYWORDS

emergency medicine; clinical decision support; health IT; human factors; work systems; SEIPS; Systems Engineering Initiative for Patient Safety; educational outreach; academic detailing; implementation method; department-based; CDS; clinical care; evidence-based; CDS tool; gerontology; geriatric; geriatrics; older adult; older adults; elder; elderly; older person; older people; preventative intervention; team-based analysis; machine learning; high-risk patient; high-risk patients; pharmaceutical; pharmaceutical sales; United States; fall-risk prediction; EHR; electronic health record; interview; ED environment; emergency department

Introduction

Background

New technologies incorporating machine learning–derived content into clinical decision support (CDS) have the potential to bring transformative improvements to clinical care [1-3]. Identifying high-risk patients who merit referral for preventative care services has historically required cumbersome screening, but now can be rapidly completed by risk prediction algorithms that consider the patient's entire electronic health record (EHR) [4-6]. By incorporating machine learning–derived content, clinicians' decision-making can be augmented by insights that may otherwise go unnoticed. Yet the potential benefits of these CDS tools will only be realized when they are designed to fit the clinical contexts in which clinicians work [3,7]. Health information technologies (HITs), including CDS tools, that fail to fit clinicians' decision-making processes and workflows are unlikely to be adopted and even risk increasing clinician burden and burnout [8-10].

However, even technologies that are designed using today's best usability guidance [11] often fail to fit the clinical context upon initial implementation [12,13]. As health systems continue to evolve in response to emergent patient needs and expectations (eg, COVID-19 and its aftermath), regulatory requirements, and staffing challenges, CDS tools are being implemented in increasingly sensitive and complex environments. While implementation science frameworks consider a variety of contextual factors [14-16] and some methods exist for assessing and identifying them [17,18], there is a gap in methods for rapidly identifying contextual factors immediately postimplementation—when it may be easiest to respond to and redesign for emergent barriers to the technology's use, safety, and effectiveness [19].

One method that has the potential to be adapted to rapidly identify contextual factors influencing the implementation of HIT is *academic detailing*. A repurposing of pharmaceutical sales representatives' tactics, academic detailing is defined as a "personal visit by a trained person to health professionals in their own settings" [20]. The goal of these personal visits is to improve care quality and patient outcomes by promoting evidence-based practice through focused clinician education [21]. As an implementation method, academic detailing can be conceptualized as a combination of 3 Expert Recommendations for Implementing Change (ERIC; also known as Evidence-based Recommendations for Implementing Change) strategies: auditing and providing feedback, conducting educational outreach visits, and practice facilitation [14]. The method's attention to the specific contexts in which clinicians make decisions—both by conducting visits in situ and by discussing barriers to and strategies for making evidence-based decisions—may present a unique opportunity to not only promote the use of a newly implemented HIT but also identify contextual factors influencing its initial implementation.

Study Objective

We propose the use of academic detailing as a method for achieving two goals in the implementation of an emergency department (ED)–based CDS tool to prevent future falls: (1) ensuring the correct understanding of the tool and its evidence base and (2) identifying contextual factors influencing the tool's initial implementation. As part of a long-term goal of assessing academic detailing for achieving these 2 aims, the objective of this study was to assess academic detailing through an analysis of the resultant qualitative data.

Methods

Study Context and Setting

This study was conducted at a large academic medical center located in the Midwestern United States. The associated ED, a level 1 trauma center, treats over 60,000 patients per year. The CDS tool being evaluated is intended to facilitate both screening for outpatient fall risk among older adults presenting to the ED and the referral to a fall prevention clinic for those patients at high risk. Our research team developed an outpatient fall risk prediction algorithm from EHR data and, in concert with our partner health system, designed and implemented a CDS tool to use the algorithm that went live in July 2020 [22,23]. In November 2020, the CDS was updated such that it enforced a "hard stop" in the clinician's workflow and required them to interact with it.

Upon arrival to the ED, all patients aged 65 years and older with an in-system primary care provider are assessed for fall risk algorithmically based on their extant EHR data. For eligible patients who are at high risk for falls, during the discharge process, an interruptive CDS alert is shown to clinicians, which informs them of the patient's risk factors and expedites the placement of a referral order to a mobility and falls clinic, an evidence-based preventive intervention. Patients who are referred are informed both by the nursing staff and in writing and are contacted to schedule an appointment by scheduling staff in the days following their ED visit. This intervention has been described in more detail elsewhere [23,24].

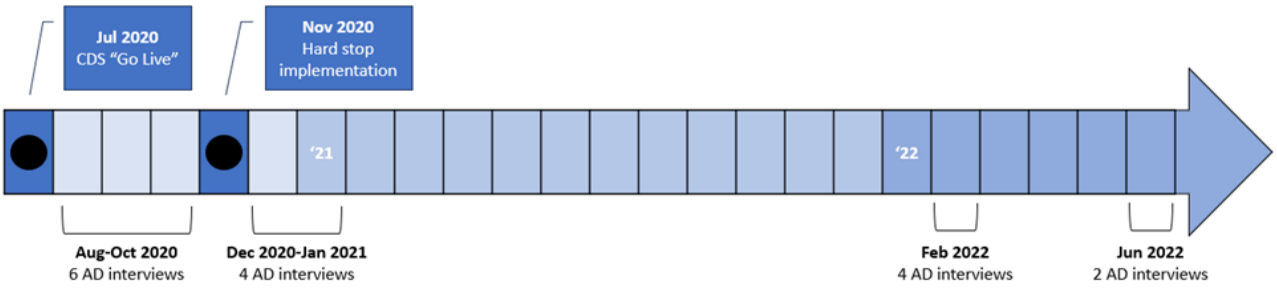
Study Design

To assess academic detailing as a method for simultaneously achieving the goals of ensuring the CDS was understood and identifying contextual factors influencing its implementation, we used a qualitative approach. We conducted 16 semistructured academic detailing interviews with emergency medicine resident

XSL•FO
RenderX

physicians (n=10) and advanced practice providers (n=6) who had previously encountered our CDS tool in practice, that is, within the last month. All interviews took place between August 2020 and June 2022, with 6 of the 16 interviews occurring prior to the implementation of the CDS hard stop (Figure 1). We purposively selected a range of participants based on how frequently they responded to the CDS. The academic detailing interviews were led by an intervention expert (AM) who had a comprehensive understanding of the CDS tool and thus was able to identify and correct any misconceptions about the tool and its use—a critical aspect of effective academic detailing [21].

Figure 1. Timeline of clinical decision support implementation and academic detailing interviews. AD: academic detailing; CDS: clinical decision support.



Study Procedure

Interviews were roughly 10 minutes long and took place over the phone or in person while the clinician was on shift. The intervention expert used an interview guide developed using the critical incident technique, which asks the participants to mentally put themselves in the moment they first saw the tool in the EHR [25]. The interview guide (Multimedia Appendix 1) contained questions such as "How and when did you see the tool initially? What was your reaction? How did you make the decision to refer the patient or not?" Additionally, comments made by a participant that suggested an incomplete or inaccurate understanding of the tool were addressed by the intervention expert (eg, "the tool refers patients to the Mobility and Falls clinic, not the Faint and Falls clinic"). Phone interviews were transcribed in real time, while notes were taken during in-person interviews and then written out immediately after.

Ethical Considerations

This study was reviewed and determined to be exempt by the UW-Madison Health Sciences IRB (ID# 2020-1100). Participants were not compensated, and data were deidentified for analysis.

Data Analysis

We conducted an inductive, team-based content analysis [26,27]. Two researchers (AM and MAL) began by independently reviewing and coding 4 interviews, line-by-line, to identify factors that influenced clinicians' decision-making. The researchers then met to compare and refine codes until there was agreement. This process continued iteratively until all interviews were coded; the resultant codebook contained 31 codes and subcodes (eg, patient risk factors and clinician communication). Another researcher (HJB) generated categories of factors that influenced clinicians' referral or nonreferral from the codes through a process of organizing similar and dissimilar codes, periodically incorporating feedback from the research team, until there was agreement.

Results

Overview

We identified five categories of factors that impacted clinicians' use of the CDS: (1) aspects of the CDS tool's design, for example, its features, usability, and how it fits in the clinician's workflow; (2) clinicians' understanding (or misunderstanding) of the CDS or referral process; (3) the busy nature of the ED environment; (4) clinicians' perceptions of the patient and their associated fall risk; and (5) the opacity of the referral process. Table 1 organizes the identified factors by these categories, including a description of the type of clinician education that was done during the academic detailing interviews.

XSL•FO
RenderX

Table 1. Factors impacting clinicians' CDS[a] use identified through academic detailing and clinician education done during academic detailing interviews. Positive and negative factors are indicated by the +, +/–, and – symbols.

Categories of factors impacting clinicians' CDS use through academic detailing	Factors impacting clinicians' CDS use identified through academic detailing	Clinician education done during academic detailing interviews
Aspects of the CDS tool's design	[+] CDS is simple. [+] CDS requires minimal input from the clinician. [+] CDS automatically identifies a high-risk patient and prompts care that the clinician would not otherwise have considered. [+/–] CDS enforces a hard stop in the clinician's workflow. [+/–] CDS alert fires while the clinician is completing discharge in the EHR[b].	• Discussed why the CDS alert fires when it does and the potential benefits and challenges of it firing at a different point in the clinician's workflow.
Clinicians' understanding of the CDS or referral process	[–] Clinician confuses the geriatric mobility and falls clinic with the faint and falls clinic. [–] Clinician believes only patients being seen for a fall are appropriate referrals. [–] Clinician believes referring the patient will be cumbersome, ie, require written justification. [+] Clinician is familiar with the concept of the CDS from an organizational stakeholder's communication.	• Clarified which clinic the referral goes to. • Clarified that referral is appropriate preventative care for patients regardless of their presenting problem. • Demonstrated how simple and quick it is to place a referral via the CDS.
Busy nature of the ED[c] environment	[+/–] A busy ED environment.	
Clinicians' perceptions of the patient and their associated fall risk	[+/–] Clinicians' agreement with the CDS's assessment of the patient's fall risk. [–] Clinicians' perception of the patient's openness to, need for, or benefit from the intervention.	• Demonstrated where in the CDS to find the reasons the patient is being flagged as high risk. • Stressed the potential benefits of a successful referral for both the patient and health system.
Opacity of the referral process	[–] Clinicians lack clarity on where the referral goes once it is sent. [–] Clinicians are uncertain about who should communicate with the patient about the referral, ie, themselves or a nurse. [+/–] Clinicians (do not) have the information necessary for counseling patients on what to expect from the referral and why they are being referred.	• Clarified which clinic the referral goes to. • Clarified the importance of counseling patients on the referral and demonstrated where in the CDS to access information to support counseling patients on the referral. • Demonstrated where in the CDS to find the reasons the patient was flagged as high risk.

[a]CDS: clinical decision support.

[b]EHR: electronic health record.

[c]ED: emergency department.

Aspects of the CDS Tool's Design

The first category of factors that influenced clinicians' use of the CDS was those that related to the design of the CDS tool. Many clinicians described the CDS as user friendly or easy to use, citing the limited number of clicks required and that the CDS did not require the clinician to enter any text. Further, clinicians found that the automatic nature of the CDS tool supported them in providing appropriate care that they otherwise would not have considered:

> *I appreciated how it fired on its own. I wasn't even thinking about falls in the patient because he came in for [condition], not a fall. When it fired, I realized he was a great candidate, but it wasn't something I thought about prior.* [Participant 12]

Further, clinicians described how the CDS integration into their workflow impacted their use of the CDS. For one, the CDS enforced a "hard stop," requiring the clinician to interact with it. While clinicians' feelings on the hard stop varied from being annoyed to finding it valuable, it was only described as impacting CDS use during high-volume times in the ED, which is discussed further in the next section. Clinicians generally appreciated where the CDS fit into their workflow—upon discharging the patient in the EHR—such that it "doesn't seem like an additional step" (participant 15). Some clinicians noted the timing of the CDS as being too late in the care process to have a personal discussion with their patient about the referral, for example, they are discharging their patient in the EHR after they have already done their final visit to the patient's room. Thus, when clinicians saw the CDS alert after their final visit with the patient, they described being less likely to refer the patient because of the additional time necessary to go back to the room to discuss the referral. However, most clinicians thought the CDS alert was well situated in their workflow:

> *It just pops up at discharge. It's just a click, the referral order is already filled out. It's very easy to use. It adds maybe 20 seconds to the discharge process.* [Participant 14]

XSL•FO
RenderX

Clinicians' Understanding of the CDS and Referral Process

The second category of factors that influenced clinicians' use of the CDS was clinicians' understanding, or misunderstanding, of the CDS and the referral process. One such misunderstanding of the CDS was clinicians confusing a separate faint and falls cardiology Clinic with the actual target of the referral, the mobility and falls clinic, which specifically addresses geriatric falls. Consequently, 2 clinicians cited their nonreferral as due to the inappropriateness of the faint and falls clinic for their geriatric patient:

> *I think more of the syncope patients and possible cardiac or peripheral vertigo patients who don't need to be admitted or are younger and are less high risk and are anxious about having syncopal episodes for the first time…. We want to get [them to] an outpatient visit so they'll be more likely to follow up with the clinic and have [care] done. I don't think of it as much who just have a mechanical fall.* [Participant 1]

Further, 6 clinicians described that they would not refer a patient who was being seen for another chief complaint, believing that only patients being seen for a fall are appropriate referrals.

Additionally, 3 clinicians expressed concerns about how cumbersome they believed the referral would be; however, their perception of the tool changed immediately once it was demonstrated that it only required 2 clicks. For example, participant 2 said:

> *The BPA would be less annoying if I knew I didn't need to justify it. If I had known that's all I had to do, I would have clicked [to accept the referral].*

One clinician stated that they expected that the CDS would be cumbersome, that is, that they would have to write out the referral because they thought "[the CDS] would be like the other referrals" that they had come across in the EHR (participant 8).

On the other hand, 1 factor in this category that positively impacted clinicians' use of the CDS was that the CDS was familiar to some clinicians given previous communication from an organizational stakeholder (ie, clinical champion). For example, participant 12 said:

> *When the CDS fired, I knew [clinical champion] sent an email about this. From my interactions with this patient, I thought [they were] at high risk of fall and knew [they would] benefit from it. That's why I tried to place that consult.*

Busy Nature of the ED Environment

The third category influencing CDS use was the busy nature of the ED environment. Five clinicians described the ED environment as a factor impacting their CDS use, whereas at least 1 clinician explicitly said, "the ED environment wouldn't affect whether or not I refer a patient" (participant 4). Clinicians varied in their description of the impact of the busy ED environment, ranging from "I would ignore the [CDS]" (participant 12) to "I just do the referral" (participant 14). Those clinicians who said that the ED environment increased their likelihood of referring the patient cited the CDS's hard stop and a significant amount of text as associated factors that shaped their decision-making. While other clinicians described the amount of text in the CDS as a stressor, it was not otherwise described as influencing clinicians' use of the CDS.

Clinicians' Perceptions of the Patient and Their Associated Fall Risk

The fourth category of factors influencing the use of the CDS was clinicians' perceptions of the patient and their associated fall risk. Overall, most clinicians (10/16) agreed with the CDS's assessment of the patient's fall risk. One clinician said:

> *In general, my reaction has been "oh that kinda makes sense." It was always kind of a surprise in the sense that I hadn't really considered the risk of falls before, but it never seems outlandish that that was a potential concern.* [Participant 3]

A few clinicians described instances of being annoyed by the firing of the CDS when it seemed irrelevant and thus did not use it. Conversely, another clinician was frustrated when the CDS did not fire when they expected to see it. Clinicians described occasionally not referring patients because they appeared to be "independent and functional" (participant 11) or "generally active and stable" (participant 10), or because their fall was "strictly mechanical" (participant 7).

Further, clinicians' perception of the patient's openness to or need for intervention impacted their CDS use. One clinician described factoring their assessment of the patient's openness to going to the mobility and falls clinic into their decision to refer the patient or not:

> *You can kind of get a vibe if someone is going to the doctor. If you tell them there's another doctor you can see; if they don't even want to talk to me, I doubt they're going to go to another doctor. If it's going to be useless, I don't want to waste everyone's time. I like to tell people and if they say I'm not going to that then I won't refer. I think I dictate if I'm going to [refer the patient] based on the conversation.* [Participant 8]

Clinicians also described being more likely to refer patients who they perceived as having a greater need for the intervention. For example, a clinician said, "If the patient seems more anxious or [they] don't have as good of a support system or advocate, I would refer them" (participant 1). Alternatively, clinicians also described assessing how the referral would fit into the patient's care plan, for example, if the patient had an upcoming surgery, the clinician would opt not to refer the patient so as to not "throw an extra thing on top of them" (participant 6).

Opacity of the Referral Process

The final category of factors influencing clinicians' use of the CDS was the opacity of the referral process. One clinician described how they lacked clarity on where the referral goes once it is sent by saying, "it feels like I'm just sending the referral off to the void and I don't know who they're getting referred to" (participant 3). In contrast to a clinician confusing the mobility and falls clinic with the faint and falls clinic,

XSL•FO
RenderX

discussed previously, this clinician was specifically pointing to the lack of feedback they received about how the process of referring a patient to the mobility and falls clinic unfolds over time. Consequently, another factor clinicians said impacted their use of the CDS was the ambiguity around who should communicate with the patient about the referral: themselves or a nurse. Clinicians described this factor as being more prominent if they had already spoken to the patient and thus referring the patient would require them to initiate another conversation with the patient themselves or "hope the nurse will tell the patient" (participant 1).

Finally, a few clinicians said they lacked the necessary information to counsel the patient—either about what to expect from the referral and the mobility and falls clinic or about the reasons the patient had been flagged as high risk for falls. One clinician suggested that having guidance on how to counsel the patient might make referring patients an easier choice:

> *I haven't been really having detailed conversation about what this entails and what they should expect. In the moment I hadn't quite seen a link on how to counsel patients on this referral…I do want to refer patients…I just wish I knew what to tell patients.* [Participant 3]

Clinicians also described lacking sufficient information to explain to patients why they were flagged by the CDS as high risk for falls. In particular, clinicians said this information would likely influence their referring of patients being seen for chief complaints other than a fall by making it easier to explain the referral to the patient.

Clinician Education Done During Academic Detailing

During the academic detailing interviews, various misconceptions were addressed directly by the intervention expert through clinician education. First, 1 misconception described previously was the mistaken belief that a referral to the mobility and falls clinic would be inappropriate for people being seen for a chief complaint other than a fall. The intervention expert addressed this by clarifying that the CDS alert fires for any older adult being seen in the ED who is at high risk of falling in the future regardless of their presenting complaint, so barring any contraindications—for example, patient in hospice—it would be appropriate to refer the patient. The intervention expert also clarified, for clinicians who misunderstood, the correct target clinic of the referral (ie, the mobility and falls clinic). Generally, the intervention expert stressed the potential benefits of a successful referral for both the patient (eg, improved quality of life) and the health system (eg, reduced use).

Another misconception that was addressed via academic detailing was the perception that referring a patient would be too cumbersome. By demonstrating that accepting the CDS alert and placing a referral takes only 2 clicks, this misconception was promptly addressed. The intervention expert also demonstrated where in the CDS to access information to support counseling patients on the referral and where in the CDS to find the reasons the patient was flagged as high risk. Finally, for any clinicians who had issues with where the CDS alert fired in their workflow, the intervention expert discussed the reasons for the alert firing when it does and the potential benefits and challenges of it firing at a different time. Oftentimes, after discussion, the clinician had a new appreciation for the complexity of designing the CDS alert.

Discussion

Principal Findings

This study demonstrates the use of academic detailing for supporting the early implementation of HIT, allowing us to identify and begin to address factors that impacted clinicians' use of the CDS while concurrently educating clinicians to ensure the correct understanding of the CDS tool and intervention. By bundling multiple ERIC strategies, academic detailing appears to be a promising method for providing timely feedback to improve HIT implementation.

Addressing Contextual Factors Within Detailing Sessions

A key component of the academic detailing method is its emphasis on clinician education [21] which, in the context of our study, involves correcting clinicians' misconceptions. For example, 1 misconception that we identified and addressed through clinician education was the mistaken belief that a referral to the mobility and falls clinic was only appropriate for people being seen for a fall. Given the nature of this CDS tool, that is, its ability to predict future risk, the impact of this misconception is that the opportunities to intervene in the routine care of high-risk patients being treated for other chief complaints would be missed. As participant 12 articulated, quoted in the "aspects of the tools design" results section, a particular value of the CDS tool is that it runs automatically, that is, does not require clinician initiation; thus, it can prompt the clinician to consider fall risk—and care to address that risk—that they may not have been considering previously. Embedding clinician education into academic detailing thus addressed a high-impact misconception with immediacy. However, it remains to be seen whether and how addressing these misconceptions translates to clinicians' use of the CDS tool. Our future work will explore the impact of these academic detailing sessions on implementation incomes, for example, clinicians' rates of referral and their acceptance of the tool.

Another important misconception to address within the academic detailing interviews was the perception that referring a patient would be too cumbersome. By demonstrating the simplicity of accepting the CDS alert and placing a referral, this misconception was promptly addressed which likely prevented its propagation. However, as described in the Results section, clinicians' perceptions of the CDS tool are situated within the context of the existing EHR and thus are beholden to a broader understanding of how similar tools work (ie, a mental model) [28]. As such, clinicians' responses to CDS alerts can be understood to be habitual, triggered by environmental cues [29]; therefore, solely addressing this misconception at the clinician level is unlikely to sustain CDS use over time. Altering clinicians' mental models of CDS tools and the EHR warrants systems-level redesign.

XSL•FO
RenderX

The content of the clinician education that is included in academic detailing is paramount to its success in increasing the use of an intervention [21]. Previous literature also notes the importance of the relationship between the clinician and the person doing the clinician education [21,30]. For this study, the intervention expert who conducted the academic detailing interviews had extensive experience working with the ED staff and had developed a rapport with them. To carry out academic detailing in another setting, there may be initial relationship- and trust building to do to achieve the detailed results our intervention expert was able to capture. Yet, given their role as a researcher (vs a fellow clinician), there were potentially missed opportunities for educating clinicians on topics that would have been better received from a colleague. For example, the deeply entrenched custom of referring to many older adults' community-based falls as being "mechanical," a catch-all term for falls that does not have an emergent, addressable cause, is known to negatively affect care [31]. This could have potentially been addressed by a colleague; however, in this study, we did not address this clinician perception as it fell outside of the expertise of our intervention expert, that is, outside of the purview of the CDS tool and the referral it recommends.

Addressing Contextual Factors via Redesign

The factors impacting clinicians' use of the CDS point directly to opportunities to intervene in and improve the CDS implementation process (Textbox 1). As discussed previously, clinician education can be done immediately, within the academic detailing interview; however, the clinician education that had to be provided within the interview can inform the redesign of a better rollout (eg, addressing what are likely to be misconceptions up front). Future rounds of academic detailing should thus result in the need for less or different clinician education from the intervention expert.

Textbox 1. Potential approaches for intervening in the health information technology implementation process to improve clinical decision support acceptance and use.

Real time (within academic detailing interview)

- Demonstrating the current capabilities and function of the tool, for example, how easy it is to place a referral, where to access information about why the patient was flagged as high risk, and information to support counseling the patient on the referral.
- Discussing why the clinical decision support tool works the way it does and the potential benefits and challenges of redesigning it.
- Clarifying how the referral works, where it goes, and who is an appropriate candidate for the intervention.
- Addressing problematic or harmful misconceptions, for example, that there is no role the emergency department can play in providing preventive care after "mechanical falls."
- Discussing how successfully using the clinical decision support and placing a referral improves patient outcomes and health system outcomes, for example, by reducing future visits to the emergency department.

Short term (quick fixes)

- Attending regularly scheduled meetings with clinicians to remind them about the clinical decision support and clarify misconceptions about placing the referral.
- Associating the organizational stakeholder's name or image with the clinical decision support.
- Adding the mobility and falls clinic information to the clinical decision support, that is, the phone number and location.

Long term (adaptation and redesign)

- Developing feedback mechanisms for clinicians to hear about successfully referred patients.
- Clarifying roles around patient communication, that is, what is communicated by the clinician versus the nurse, and designing the clinical decision support to support those roles.
- Reviewing clinical decision support tools for potential interaction effects, for example, 2 clinical decision support tools fire on similar populations and are likely to be confused.
- Providing talking points on what the patient can expect after discharge with respect to scheduling and going to an appointment with the mobility and falls clinic.
- Providing talking points that explain why patients being seen for issues other than falls may be referred.
- Personalizing the timing of the clinical decision support alert for clinicians who tend to talk to patients before completing the discharge in the electronic health record, for example, moving the clinical decision support alert earlier in clinicians' workflow.

In the longer term, a variety of approaches could be used to address the factors we identified as impacting the clinicians' use of the CDS. For one, reviewing the CDS tools that are currently implemented in overlapping clinical contexts could identify potential interactions with the newly implemented CDS. To avoid interaction effects, the new CDS could be redesigned to differentiate it from others, for example, to alert a more specific patient population or to have clear visual cues and messaging. Alternatively, a review of the CDS ecosystem may prompt the removal of underused or ineffective CDS tools. Recent research, while limited, suggests that health systems that optimize CDS alerts, that is, reduce unnecessary or less useful alerts, see improved CDS use [32]. Further, those effects

XSL•FO
RenderX

are not limited to the optimized CDS but spread to other CDS in the system [32].

Other redesigns that would address factors identified through academic detailing could address the workflow integration of the CDS. For example, for the clinicians who typically talk to a patient before completing the discharge in the EHR, moving the firing of the CDS alert earlier in the clinical workflow may be warranted. Beyond considering the timing of the CDS, to achieve workflow integration as defined by Salwei et al [33,34], the design of the CDS should consider the dimensions of flow, scope of patient journey, and level. An example of such a redesign could be—in the case where the fall risk CDS alert would happen earlier in the clinician's workflow—allowing the clinician to "snooze" the alert until the point at which they have discussed the mobility and falls clinic referral with the patient. This design would increase the chance that the clinician would see the CDS prior to speaking with the patient for the last time, which could promote more meaningful patient counseling on fall risk; however, this design could also have unintended consequences, which should be explored prior to broad implementation.

In designing for CDS use, it is important to remember that increased use does not always equate to increased *appropriate* use (ie, referrals for patients that are a good fit for the mobility and falls clinic intervention). Thus, the findings from academic detailing should also be considered in light of, and be used to design to support, successful teaming between the CDS tool and the clinician. A potential design to promote teamwork between the CDS tool and the clinician could be to include on the CDS a list of exclusion criteria for the mobility and falls clinic that the CDS tool is unable or poorly able to assess (eg, late-stage dementia). The clinician, then, when considering referring the patient would be alerted to where their clinical judgment is especially necessary.

Work Systems Approach to Redesign

Given the breadth of potential redesign options and the challenge of prioritizing efforts to improve not only CDS use, but *appropriate* CDS use, it is pertinent to consider models that can hold and make sense of system complexity. One model that has proven to be valuable across a variety of health care domains and in supporting the design of technologies—the Systems Engineering Initiative for Patient Safety (SEIPS) model—conceptualizes the work of clinicians as happening in a *work system*, which invariably influences care processes and outcomes [35-37]. The SEIPS model, which synthesizes literature on job stress, job design, and health care quality [37,38], provides a theoretical foundation for understanding why the system is achieving certain outcomes and how the system may be redesigned to achieve alternative outcomes.

In a parallel analysis—presented elsewhere [39]—we found that the data we collected using the academic detailing method successfully mapped to the SEIPS model's work system components, for example, the *people* who do the work, the *tasks* they complete, the *tools and technologies* they use, and the *physical* and *organizational environment* they work in. A key aspect of the SEIPS model is the conceptualization of balance—that work system components that negatively influence processes and outcomes (barriers) may be balanced by positive components (facilitators) [37,40]. Thus, through redesign efforts, we can either seek to address the work system barrier or enhance the work system facilitator. Applying a work systems approach to system redesign to address the factors we identified through academic detailing has the potential to result in more sustainable HIT implementation.

Beyond redesigning the CDS itself, as discussed in the previous section, redesigning the work system to clarify the process of referring a patient to the mobility and falls clinic may be essential to promoting the appropriate use of the CDS. This would require creating clarity around who should communicate and about what with the patient (ie, the referring clinician and the nurse). Further, creating transparency around the positive outcomes of past referrals to the mobility and falls clinic (ie, success stories) may promote trust in the referring clinicians that this is an action worth taking.

Limitations

The following limitations of our study should be considered. First, the academic detailing method, as applied here, relies on the clinicians to report what they perceive as influencing their use of the CDS. However, it is possible that clinicians' perceptions differ from what they actually do—a common challenge in understanding people's work is the difference between "work as imagined" versus "work as done" [41]. Second, this study focuses on academic detailing around a specific CDS tool that produces an interruptive alert to which a clinician must respond that they agree or decline to refer the patient. It is possible that there are other considerations for CDS tools and HIT that operate differently from this study (eg, tools that require more in-depth information processing or that must be initiated by the clinician). Further, given this academic detailing method was applied in a live ED setting over nearly 2 years—including multiple waves of COVID-19—a variety of external factors may have contributed to clinicians' use (and perception of their use) of the CDS. Finally, it is yet unclear how many rounds of academic detailing would be required to capture and address the majority of factors impacting the implementation of the HIT. Future research should explore the use of the academic detailing method over a broader range of the implementation process so that the effort and resources required to conduct the interviews are used most effectively.

Conclusions

With HIT developing at rapid speeds, it is essential we develop methods to support its integration into the complex environments in which they will be used. From our initial study, it appears that academic detailing is a promising method for both promoting the correct understanding of a CDS tool and identifying contextual factors influencing its implementation. Thus, academic detailing can inform real-time adjustments of a tool's implementation (eg, refinement of the language used to introduce the tool), and larger scale redesign of the CDS tool to better fit the dynamic work environment of clinicians.

XSL•FO
RenderX

Acknowledgments

The authors would like to acknowledge the clinicians who participated in our academic detailing interviews and shared their insights with us. This work was supported by the Agency for Healthcare Research and Quality (grant R18HS027735; BWP, principal investigator). This work is solely the responsibility of the authors and does not represent the official views of the Agency for Healthcare Research and Quality.

Authors' Contributions

HJB performed data analysis and led the writing and revising of the manuscript. AM performed data collection and analysis and critically reviewed and revised the manuscript. MAL performed data analysis and critically reviewed and revised the manuscript. DJH, DAW, and MNS critically reviewed and revised the manuscript. BWP conceptualized and oversaw the study and critically reviewed and revised the manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Interview guide for academic detailing interviews.
[DOCX File , 22 KB - humanfactors_v11i1e52592_app1.docx]

References

1. Emanuel EJ, Wachter RM. Artificial intelligence in health care: will the value match the hype? JAMA 2019;321(23):2281-2282. [doi: 10.1001/jama.2019.4914] [Medline: 31107500]
2. Obermeyer Z, Emanuel EJ. Predicting the future—big data, machine learning, and clinical medicine. N Engl J Med 2016;375(13):1216-1219 [FREE Full text] [doi: 10.1056/NEJMp1606181] [Medline: 27682033]
3. Chen JH, Asch SM. Machine learning and prediction in medicine—beyond the peak of inflated expectations. N Engl J Med 2017;376(26):2507-2509 [FREE Full text] [doi: 10.1056/NEJMp1702071] [Medline: 28657867]
4. Carpenter CR, Griffey RT, Stark S, Coopersmith CM, Gage BF. Physician and nurse acceptance of technicians to screen for geriatric syndromes in the emergency department. West J Emerg Med 2011;12(4):489-495 [FREE Full text] [doi: 10.5811/westjem.2011.1.1962] [Medline: 22224145]
5. Donaldson MG, Khan KM, Davis JC, Salter AE, Buchanan J, McKnight D, et al. Emergency department fall-related presentations do not trigger fall risk assessment: a gap in care of high-risk outpatient fallers. Arch Gerontol Geriatr 2005;41(3):311-317. [doi: 10.1016/j.archger.2005.04.008] [Medline: 15982758]
6. Goldstein BA, Navar AM, Pencina MJ, Ioannidis JPA. Opportunities and challenges in developing risk prediction models with electronic health records data: a systematic review. J Am Med Inform Assoc 2017;24(1):198-208 [FREE Full text] [doi: 10.1093/jamia/ocw042] [Medline: 27189013]
7. Kawamoto K, Houlihan CA, Balas EA, Lobach DF. Improving clinical practice using clinical decision support systems: a systematic review of trials to identify features critical to success. BMJ 2005;330(7494):765 [FREE Full text] [doi: 10.1136/bmj.38398.500764.8F] [Medline: 15767266]
8. Carayon P, Hoonakker P, Hundt AS, Salwei M, Wiegmann D, Brown RL, et al. Application of human factors to improve usability of clinical decision support for diagnostic decision-making: a scenario-based simulation study. BMJ Qual Saf 2020;29(4):329-340 [FREE Full text] [doi: 10.1136/bmjqs-2019-009857] [Medline: 31776197]
9. Hoonakker PLT, Carayon P, Salwei ME, Hundt AS, Wiegmann D, Kleinschmidt P, et al. The design of PE Dx, a CDS to support pulmonary embolism diagnosis in the ED. Stud Health Technol Inform 2019;265:134-140 [FREE Full text] [doi: 10.3233/SHTI190152] [Medline: 31431589]
10. National Academies of Sciences, Engineering, and Medicine, National Academy of Medicine, Committee on Systems Approaches to Improve Patient Care by Supporting Clinician Well-Being. Taking Action Against Clinician Burnout: A Systems Approach to Professional Well-Being. Washington, DC: National Academies Press; 2019.
11. Zhang J, Johnson TR, Patel VL, Paige DL, Kubose T. Using usability heuristics to evaluate patient safety of medical devices. J Biomed Inform 2003;36(1-2):23-30 [FREE Full text] [doi: 10.1016/s1532-0464(03)00060-1] [Medline: 14552844]
12. Patterson BW, Pulia MS, Ravi S, Hoonakker PLT, Hundt AS, Wiegmann D, et al. Scope and influence of electronic health record-integrated clinical decision support in the emergency department: a systematic review. Ann Emerg Med 2019;74(2):285-296 [FREE Full text] [doi: 10.1016/j.annemergmed.2018.10.034] [Medline: 30611639]
13. Greenes RA, Bates DW, Kawamoto K, Middleton B, Osheroff J, Shahar Y. Clinical decision support models and frameworks: seeking to address research issues underlying implementation successes and failures. J Biomed Inform 2018;78:134-143 [FREE Full text] [doi: 10.1016/j.jbi.2017.12.005] [Medline: 29246790]

XSL•FO
RenderX

14. Powell BJ, Waltz TJ, Chinman MJ, Damschroder LJ, Smith JL, Matthieu MM, et al. A refined compilation of implementation strategies: results from the Expert Recommendations for Implementing Change (ERIC) project. Implement Sci 2015;10:21 [FREE Full text] [doi: 10.1186/s13012-015-0209-1] [Medline: 25889199]
15. Damschroder LJ, Reardon CM, Widerquist MAO, Lowery J. The updated consolidated framework for implementation research based on user feedback. Implement Sci 2022;17(1):75 [FREE Full text] [doi: 10.1186/s13012-022-01245-0] [Medline: 36309746]
16. Nilsen P, Bernhardsson S. Context matters in implementation science: a scoping review of determinant frameworks that describe contextual determinants for implementation outcomes. BMC Health Serv Res 2019;19(1):189 [FREE Full text] [doi: 10.1186/s12913-019-4015-3] [Medline: 30909897]
17. McCormack B, McCarthy G, Wright J, Slater P, Coffey A. Development and testing of the Context Assessment Index (CAI). Worldviews Evid Based Nurs 2009;6(1):27-35. [doi: 10.1111/j.1741-6787.2008.00130.x] [Medline: 19207560]
18. Estabrooks CA, Squires JE, Cummings GG, Birdsell JM, Norton PG. Development and assessment of the Alberta context tool. BMC Health Serv Res 2009;9:234 [FREE Full text] [doi: 10.1186/1472-6963-9-234] [Medline: 20003531]
19. Tyre MJ, Orlikowski WJ. Windows of opportunity: temporal patterns of technological adaptation in organizations. Organ Sci 1994;5(1):98-118. [doi: 10.1287/orsc.5.1.98]
20. O'Brien MA, Rogers S, Jamtvedt G, Oxman AD, Odgaard-Jensen J, Kristoffersen DT, et al. Educational outreach visits: effects on professional practice and health care outcomes. Cochrane Database Syst Rev 2007;2007(4):CD000409 [FREE Full text] [doi: 10.1002/14651858.CD000409.pub2] [Medline: 17943742]
21. Yeh JS, Van Hoof TJ, Fischer MA. Key features of academic detailing: development of an expert consensus using the Delphi method. Am Health Drug Benefits 2016;9(1):42-50 [FREE Full text] [Medline: 27066195]
22. Patterson BW, Engstrom CJ, Sah V, Smith MA, Mendonça EA, Pulia M, et al. Training and interpreting machine learning algorithms to evaluate fall risk after emergency department visits. Med Care 2019;57(7):560-566 [FREE Full text] [doi: 10.1097/MLR.0000000000001140] [Medline: 31157707]
23. Jacobsohn GC, Leaf M, Liao F, Maru AP, Engstrom CJ, Salwei ME, et al. Collaborative design and implementation of a clinical decision support system for automated fall-risk identification and referrals in emergency departments. Healthc (Amst) 2022;10(1):100598 [FREE Full text] [doi: 10.1016/j.hjdsi.2021.100598] [Medline: 34923354]
24. Hekman DJ, Cochran AL, Maru AP, Barton HJ, Shah MN, Wiegmann D, et al. Effectiveness of an emergency department-based machine learning clinical decision support tool to prevent outpatient falls among older adults: protocol for a quasi-experimental study. JMIR Res Protoc 2023;12:e48128 [FREE Full text] [doi: 10.2196/48128] [Medline: 37535416]
25. Flanagan JC. The critical incident technique. Psychol Bull 1954;51(4):327-358. [doi: 10.1037/h0061470] [Medline: 13177800]
26. Hsieh HF, Shannon SE. Three approaches to qualitative content analysis. Qual Health Res 2005;15(9):1277-1288. [doi: 10.1177/1049732305276687] [Medline: 16204405]
27. Braun V, Clarke V. Using thematic analysis in psychology. Qual Res Psychol 2006;3(2):77-101. [doi: 10.1191/1478088706qp063oa]
28. Holtrop JS, Scherer LD, Matlock DD, Glasgow RE, Green LA. The importance of mental models in implementation science. Front Public Health 2021;9:680316 [FREE Full text] [doi: 10.3389/fpubh.2021.680316] [Medline: 34295871]
29. Baysari MT, Tariq A, Day RO, Westbrook JI. Alert override as a habitual behavior—a new perspective on a persistent problem. J Am Med Inform Assoc 2017;24(2):409-412 [FREE Full text] [doi: 10.1093/jamia/ocw072] [Medline: 27274015]
30. Van Hoof TJ, Harrison LG, Miller NE, Pappas MS, Fischer MA. Characteristics of academic detailing: results of a literature review. Am Health Drug Benefits 2015;8(8):414-422 [FREE Full text] [Medline: 26702333]
31. Nagaraj G, Hullick C, Arendts G, Burkett E, Hill KD, Carpenter CR. Avoiding anchoring bias by moving beyond 'mechanical falls' in geriatric emergency medicine. Emerg Med Australas 2018;30(6):843-850. [doi: 10.1111/1742-6723.13129] [Medline: 30091183]
32. Chaparro JD, Beus JM, Dziorny AC, Hagedorn PA, Hernandez S, Kandaswamy S, et al. Clinical decision support stewardship: best practices and techniques to monitor and improve interruptive alerts. Appl Clin Inform 2022;13(3):560-568 [FREE Full text] [doi: 10.1055/s-0042-1748856] [Medline: 35613913]
33. Salwei ME, Carayon P, Hoonakker PLT, Hundt AS, Wiegmann D, Pulia M, et al. Workflow integration analysis of a human factors-based clinical decision support in the emergency department. Appl Ergon 2021;97:103498 [FREE Full text] [doi: 10.1016/j.apergo.2021.103498] [Medline: 34182430]
34. Salwei ME. Human-Centered Design of Health Information Technology for Workflow Integration. Madison: The University of Wisconsin; 2020.
35. Carayon P, Hundt AS, Karsh BT, Gurses AP, Alvarado CJ, Smith M, et al. Work system design for patient safety: the SEIPS model. Qual Saf Health Care 2006;15(Suppl 1):i50-i58 [FREE Full text] [doi: 10.1136/qshc.2005.015842] [Medline: 17142610]
36. Holden RJ, Carayon P, Gurses AP, Hoonakker P, Hundt AS, Ozok AA, et al. SEIPS 2.0: a human factors framework for studying and improving the work of healthcare professionals and patients. Ergonomics 2013;56(11):1669-1686 [FREE Full text] [doi: 10.1080/00140139.2013.838643] [Medline: 24088063]

XSL•FO
RenderX

37. Smith MJ, Sainfort PC. A balance theory of job design for stress reduction. Int J Ind Ergon 1989;4(1):67-79. [doi: 10.1016/0169-8141(89)90051-6]
38. Donabedian A. The quality of care. How can it be assessed? JAMA 1988;260(12):1743-1748. [doi: 10.1001/jama.260.12.1743] [Medline: 3045356]
39. Barton HJ, Maru A, Leaf MA, Hekman D, Wiegmann DA, Shah MN. Academic detailing of emergency medicine clinicians to inform post-implementation continuous design of an algorithm-based clinical decision support tool to prevent falls. 2023 Presented at: 14th Organizational Design and Management Conference; 2023; Bordeaux, France URL: https://www.odam2023.org/_files/ugd/2a1c0b_27519dfeb37d45bd82144b59db223bee.pdf
40. Carayon P, Gurses AP, Hundt AS, Ayoub P, Alvarado CJ. Performance obstacles and facilitators of healthcare providers. In: Korunka C, Hoffmann P, editors. Change and Quality in Human Service Work. Munich, Germany: Hampp Publishers; 2005:257-276.
41. Hollnagel E, Clay-Williams R. Work-as-imagined and work-as-done. In: Rapport F, Clay-Williams R, Braithwaite J, editors. Implementation Science: The Key Concepts. London, UK: Routledge; 2022:175-177.

Abbreviations

CDS: clinical decision support
ED: emergency department
EHR: electronic health record
ERIC: Expert Recommendations for Implementing Change, also known as Evidence-based Recommendations for Implementing Change
HFE: human factors engineering
HIT: health information technology
SEIPS: systems engineering initiative for patient safety

Edited by A Kushniruk; submitted 11.09.23; peer-reviewed by K Liu, E Kerns; comments to author 30.12.23; revised version received 08.02.24; accepted 02.03.24; published 18.04.24.

Please cite as:
Barton HJ, Maru A, Leaf MA, Hekman DJ, Wiegmann DA, Shah MN, Patterson BW
Academic Detailing as a Health Information Technology Implementation Method: Supporting the Design and Implementation of an Emergency Department–Based Clinical Decision Support Tool to Prevent Future Falls
JMIR Hum Factors 2024;11:e52592
URL: https://humanfactors.jmir.org/2024/1/e52592
doi:10.2196/52592
PMID:38635318

©Hanna J Barton, Apoorva Maru, Margaret A Leaf, Daniel J Hekman, Douglas A Wiegmann, Manish N Shah, Brian W Patterson. Originally published in JMIR Human Factors (https://humanfactors.jmir.org), 18.04.2024. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on https://humanfactors.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Electronic Immunization Registry in Rwanda: Qualitative Study of Health Worker Experiences

Thaoussi Uwera[1*], BSc, MSc; Mahima Venkateswaran[2*], MBBS, PhD; Kiran Bhutada[3*], BS, MD; Eleni Papadopoulou[4*], PhD; Enock Rukundo[1*], BSc, MSc; David K Tumusiime[1*], BSc, MSc, PhD; J Frederik Frøen[2*], MD, PhD

[1]Centre of Excellence in Biomedical Engineering and eHealth, University of Rwanda, Kigali, Rwanda
[2]Centre for Intervention Science for Maternal and Child Health (CISMAC), University of Bergen, Bergen, Norway
[3]Global Health Center, Albert Einstein College of Medicine, Bronx, NY, United States
[4]Global Health Cluster, Division for Health Services, Norwegian Institute of Public Health, Oslo, Norway
[*]all authors contributed equally

Corresponding Author:
Thaoussi Uwera, BSc, MSc
Centre of Excellence in Biomedical Engineering and eHealth
University of Rwanda
KG 11 Ave Gasabo
Kigali
Rwanda
Phone: 250 788229321
Email: uwera04@gmail.com

Abstract

Background: Monitoring childhood immunization programs is essential for health systems. Despite the introduction of an electronic immunization registry called e-Tracker in Rwanda, challenges such as lacking population denominators persist, leading to implausible reports of coverage rates of more than 100%.

Objective: This study aimed to assess the extent to which the immunization e-Tracker responds to stakeholders' needs and identify key areas for improvement.

Methods: In-depth interviews were conducted with all levels of e-Tracker users including immunization nurses, data managers, and supervisors from health facilities in 5 districts of Rwanda. We used an interview guide based on the constructs of the Human, Organization, and Technology–Fit (HOT-Fit) framework, and we analyzed and summarized our findings using the framework.

Results: Immunization nurses reported using the e-Tracker as a secondary data entry tool in addition to paper-based forms, which resulted in considerable dissatisfaction among nurses. While users acknowledged the potential of a digital tool compared to paper-based systems, they also reported the need for improvement of functionalities to support their work, such as digital client appointment lists, lists of defaulters, search and register functions, automated monthly reports, and linkages to birth notifications and the national identity system.

Conclusions: Reducing dual documentation for users can improve e-Tracker use and user satisfaction. Our findings can help identify additional digital health interventions to support and strengthen the health information system for the immunization program.

(JMIR Hum Factors 2024;11:e53071) doi:10.2196/53071

KEYWORDS

childhood immunization; electronic immunization registry; digital health interventions

Introduction

In 2021, a reported 18.2 million infants worldwide did not receive basic immunization, and an additional 6.8 million were only partially vaccinated, with associated higher deaths in low- and middle-income countries (LMICs) [1,2]. Health systems worldwide are adopting digital tools to improve immunization service provision as well as monitoring [3]. Digital health interventions (DHIs) have the potential to improve the management and use of health information to enhance health worker performance and provision of care and ultimately improve health outcomes [4,5]. In LMICs, electronic

XSL•FO
RenderX

immunization registries (EIRs) were initiated to support improved vaccination coverage among children, primarily through better tracking of children by combining vaccine information from different sources into a single digital record [6,7].

DHIs (in the form of EIRs) are important for immunization programs. For clients, they can help to remind families through SMS text messaging when immunization is due or has been missed. For health workers, they can help ensure that children get the vaccinations they need, improve and simplify the reporting of immunization data, identify high-risk populations for targeted interventions, and allocate resources efficiently and effectively [5-7]. EIRs, enhanced by data-driven DHIs, can help the immunization program achieve its goals of effective immunization coverage and real-time data for decision-making. EIRs can serve their purpose for immunization programs even better if integrated and synergized with DHIs for other programs such as Civil Registration and Vital Statistics (CRVS) and the national identification system. For instance, registration of all newborn babies in EIRs can improve tracking of immunization status and monitoring coverage [7]. EIRs integrated with other programs can strengthen other health services for children by providing a database of newborn babies in the population. Examples include newborn metabolic screening and childhood nutrition programs for the identification and referral of malnourished children [7].

Despite the many opportunities, several challenges hinder the effectiveness of EIRs in LMICs, such as the increased burden of data collection for health workers, which is the result of maintaining paper and digital documentation and reporting systems [7]. The implementation of EIRs, similar to all DHIs, should be aligned with the needs, both in terms of addressing the concerns of the intended users and being relevant to the users [8]. However, there is limited evidence on how to implement digital tools most effectively and sustainably across the full range of health systems [9]. The World Health Organization has highlighted the need for implementation research to identify the crucial factors that affect the implementation of DHIs for health system strengthening [5]. Implementation research can provide a systematic understanding of users' perceptions and experiences and thus enhance the usability and acceptability of DHIs.

In Rwanda, children from 0 to 15 months of age are provided with vaccines against 11 infections according to the Expanded Program on Immunization (EPI), namely, tuberculosis, poliomyelitis, diphtheria, tetanus, measles, pertussis, hepatitis B, *Haemophilus influenzae* type B, rubella, *Streptococcus pneumoniae*, and rotavirus [10]. The latest report from the Global Alliance for Vaccines and Immunization from 2017 identified issues with the immunization health information system such as data quality, population denominators based on projections from census data, and implausible coverage rates of more than 100% [11], similar to other contexts in eastern and southern Africa [12]. Incidents of vaccine dropouts and incomplete immunization, particularly for Pentavalent 3, were also identified. Significant geographic variations in immunization rates were reported, with 1 district in the northern part of the country reporting an overall coverage rate of as low as 88% [11,13].

The introduction of an EIR, known as e-Tracker, was launched in 2019 with the goal of improving overall data quality, data availability for monitoring of immunization defaulters or dropouts, and ultimately increasing immunization coverage [14]. The newly implemented e-Tracker has not yet been subject to research-based evaluations. The aim of this study was to assess the extent to which the immunization e-Tracker responds to stakeholders' needs and identify key areas for improvement in Rwanda's childhood immunization program.

Methods

Study Setting

This study was conducted among immunization nurses and data managers. Supervisors were included at the district hospital level. Health facilities were randomly selected from 5 districts in Rwanda—Gasabo, Rwamagana, Kamonyi, Gicumbi, and Rubavu, 1 from each of the 4 provinces and the City of Kigali of Rwanda. Gicumbi district, which is in the north of Rwanda, has 16 health centers; Kamonyi, in the south, has 13 centers; Rwamagana, in the east, has 15 health centers; Rubavu, in the west, has 13 centers; and Gasabo, in the central city of Kigali, has 16 health centers. In Kamonyi district and Gicumbi district, the routine immunization coverage rates for Pentavalent 3 and measles-rubella 1 in 2018 were 84% and 85%, respectively, while the coverage rate was higher than 89% in the remaining 3 districts. Gicumbi, Gasabo, and Kamonyi were among the districts with the largest percentage of underimmunized children, especially for the third dose of Pentavalent. The e-Tracker was introduced and operationalized in health centers in all districts of Rwanda in 2019.

The study participants were primary users of the immunization e-Tracker, either entering data or using the data: immunization nurses, data managers, and EPI supervisors.

Immunization-related services are organized at different levels of the health system. At the village level, community health workers engage with residents to raise awareness about childhood immunization. All primary health care services, including childhood immunization, are decentralized to the health center level. Immunization is provided at the health centers by immunization nurses or at health posts by the same nurses through community outreach in hard-to-reach areas. There are 499 health centers and 476 health posts in Rwanda [15]. More than 90% of children are immunized at the health center. All immunization sites (centers and posts) have weekly schedules of immunization days.

A health center typically has 2 immunization health workers, a nurse in charge of immunizations, and an assistant to deliver vaccines and keep records of all information pertaining to immunizations.

Figures 1 and 2 show the workflow of immunization at the health facility and the e-Tracker registration process and data visualization, respectively.

XSL•FO
RenderX

Figure 1. Workflow of immunization pertaining all the duties of a health worker (immunization nurse) on a vaccination day.

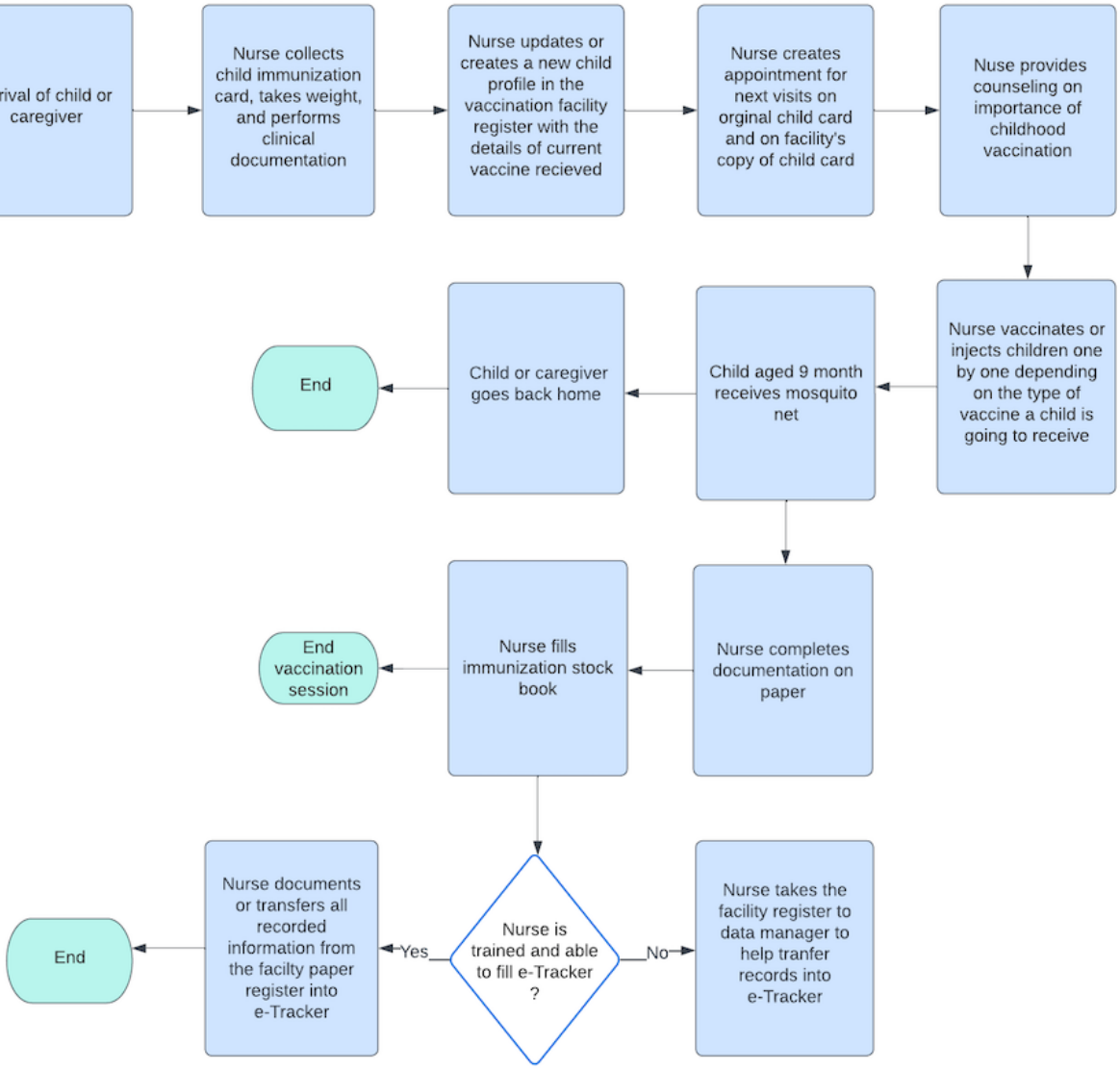
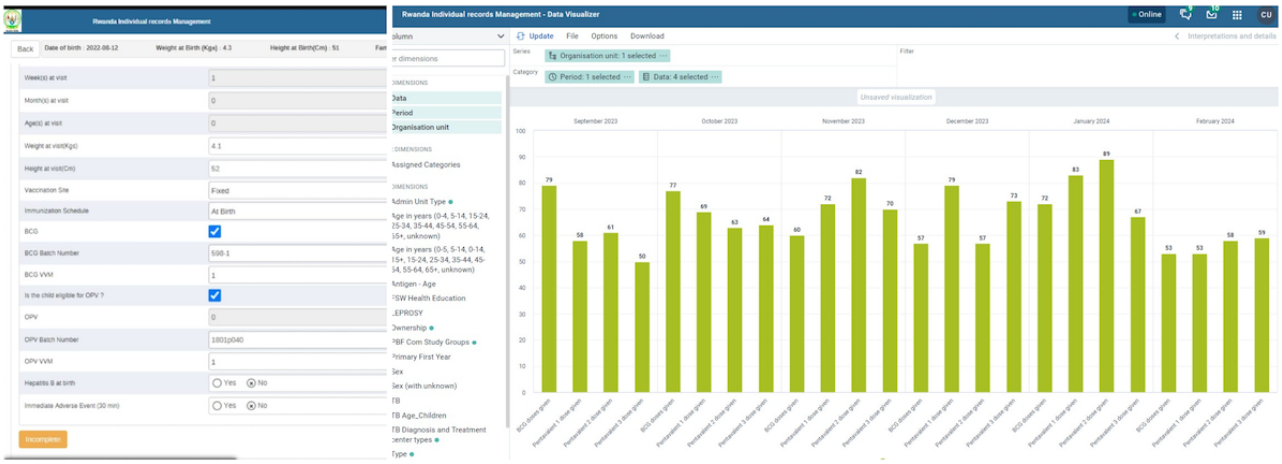


Figure 2. Immunization e-Tracker showing the client registration page and data visualization dashboard.



e-Tracker Implementation and Use

Implementation of the e-Tracker started in 2019 and was operational in all public health facilities. The e-Tracker runs on the District Health Information Software 2 (University of Oslo) platform, one of the most widely used digital health information systems globally [16]. Three cadres of health workers were trained to use the e-Tracker—immunization nurses, data managers, and EPI supervisors. All individual information are first recorded on 2 sets of paper-based forms: the child's immunization cards and the health center's immunization paper registers. The immunization nurse or the data manager then

XSL•FO
RenderX

transfers the same information from the paper registers to the e-Tracker. At the end of the month, a set of predefined data is aggregated onto paper reporting forms by immunization nurses and handed over to the data manager, who then enters these data into the aggregate reporting system built in a separate instance of District Health Information Software 2, as a part of the health management information system (HMIS).

EPI supervisors, located at the district hospitals, use the e-Tracker to assess the progress of health facilities by comparing the number of children registered as successfully vaccinated on each indicator against the monthly target provided to the health center (Table 1). The target is an estimate population based on the expected number of births in the area based on census data.

Users who have technical issues with e-Tracker can contact the central help desk. Phone calls or WhatsApp groups are typically used to resolve simple technical issues such as password reset, and for complicated issues, through visits to health centers. Immunization nurses from the health centers have a joint WhatsApp group with their respective EPI supervisors where they communicate issues regarding immunization and e-Tracker–related technical support in their district.

Table 1. Intended use and user roles in the immunization e-Tracker.

User	Intended use and user roles in the e-Tracker
Immunization nurse[a]	• Data entry and registration of new children for immunization • Update and follow up on subsequent immunizations until a child has completed his or her vaccination calendar
Data manager[a]	• Data entry and registration of new children for immunization • Update and follow up on subsequent immunizations until a child has completed his or her vaccination calendar • Generate reports of comparisons of the health center's immunization coverage rate against the target
EPI[b] supervisor	• Review reports from all health centers in the district catchment area and provide recommendations and feedback for improvement based on the data

[a]Data entry tasks could be shared by immunization nurses and data managers.

[b]EPI: Expanded Program on Immunization.

Study Design and Sampling

This study is an implementation research design that used descriptive qualitative methods [17,18] and formative evaluation to assess the extent to which the immunization e-Tracker responds to stakeholders' needs and identify key areas for improvement [19]. This was done through key informant interviews. The Human, Organization, and Technology–Fit (HOT-Fit) evaluation framework guided the data collection and analysis [20]. We chose the HOT-Fit framework because it has the potential to evaluate health information systems; encompasses comprehensive dimensions; and measures the fit between technological, human, and organizational aspects, all of which are critical for system adoption [20].

To select a sample of districts, we first assessed data reports retrieved from the e-Tracker and the national HMIS in the first 3 months of 2020 for all 30 districts in Rwanda. Four immunization indicators—Bacille Calmette-Guérin (BCG) and Pentavalent (penta) first, second, and third doses (penta1, penta2, and penta3)—were reviewed by a program manager together with a researcher (TU) to calculate completeness of data in the e-Tracker (e-Tracker–reported indicator and HMIS-reported indicator). We then selected 5 districts as follows: 1 district among the best performers (Rwamagana district >80%), 1 from the worst performers (Rubavu district <15%), and 3 districts that were in the middle (Gasabo, Gicumbi, and Kamonyi districts: 50%-60%). We randomly included 6 health centers from each of the 5 districts. From this pool of health centers, key informants and participants were purposively sampled among primary users of the e-Tracker. To cover the variation of sites across the districts appropriately, we recruited 1 nurse and 1 data manager from 1 district before moving to the next to diversify the data collected.

Data Collection

This study was carried out in accordance with COREQ (Consolidated Criteria for Reporting Qualitative Research) guidelines (Multimedia Appendix 1) [21].

Based on the 3 constructs of the HOT-Fit framework, we created study-specific definitions for each of the constructs (Table 2) and formulated an interview guide with open-ended questions (Multimedia Appendix 2). Three pilot key informant interviews were conducted with immunization nurses and data managers in 1 district (Gasabo) to validate the tool prior to data collection. We further refined the questions in the guide based on the findings from these interviews (Multimedia Appendix 2).

XSL•FO
RenderX

Table 2. Specific domains of evaluation of the e-Tracker based on the constructs of the HOT-Fit[a] framework.

HOT-Fit constructs and definitions	Study constructs and definitions

Technology: Meets the need of the projected users, is convenient and easy to use, and fits the work patterns of the professionals for whom it is intended and the overall health system	• System quality
	  • Associated with system performance: ease of use, ease of learning, response time, usefulness, system flexibility, and security
	• Information quality
	  • User perspectives and quantitative data: completeness, availability, accuracy, reliability, timeliness, relevance, and consistency
	• Service quality
	  • Service delivered: technical support, quick responsiveness, assurance, empathy, and follow-up service

Human: The person who uses and the use of information output such as reports	• System use
	  • Concerned with the frequency and breadth of health information system inquiries and functions: system users, their levels of use, training, knowledge, belief, expectation, acceptance, or resistance
	• User satisfaction
	  • Evaluation of users' experience in using the system and the potential impact of the system: perceived usefulness, enjoyment, overall satisfaction and satisfaction with specific functions, and decision-making satisfaction

Organization: Nature and factors of a health care institution	• Structure
	  • Nature (type and size), management and communication, clinical process, and workflow process. Leadership, top management support, etc
	• Environment
	  • Financial source, government, politics, and type of population being served

Net benefits	• Quality of care, clinical impact, impact on patient care and communication, and facilitation of information access

[a]HOT-Fit: Human, Organization, and Technology–Fit.

Separate interview guides were used for each category of participants. The interview questions were formulated based on each user's role both in the immunization program and the e-Tracker system. For instance, we asked questions related to user-specific employment and how e-Tracker is related to his or her job. Some e-Tracker technical questions were similar such as whether e-Tracker was easy to use, easy to learn, or about how e-Tracker responds (response time). One author (TU), a current PhD candidate, with experience in IT conducted 14 in-person, in-depth interviews with key informants (e-Tracker end users in primary health care centers and EPI supervisors in their affiliated district hospitals). Interviews were conducted with only the key informant and the interviewer present. The interviews were conducted in Kinyarwanda, took place over approximately 1 hour, and were audio recorded. No notes were taken. The audio was then transcribed in Kinyarwanda and translated into English by a bilingual professional. A group of 2 researchers (TU and ER) reviewed the translations for accuracy. The study team met on a weekly basis to evaluate the data collection process. After 4 interviews per key informant category, the data collector began hearing information repetition. The research team advised undertaking 1 more interview per participant category to ensure that no new findings were discovered. Data saturation was confirmed, and data collection was stopped. No repeat interviews were carried out.

Data Analysis

Translated interview transcripts were uploaded into NVivo 12 (Lumivero). Based on the HOT-Fit framework, a codebook was developed by the team through discussion. Using this agreed-upon codebook, 2 researchers (TU and ER) individually coded the data. A deductive coding style was applied to our data. Discrepancies in coding were discussed and resolved by the team.

The HOT-Fit Framework

After coding was completed by both researchers, the team compiled the relevant data extracts. We performed a framework analysis and worked together to place the extracted data within the HOT-Fit framework [20,22]. We analyzed interview transcripts to find all possible codes from all participants. We identified and summarized codes in accordance with constructs of the HOT-Fit framework and study-specific domains (Table 2). NVivo 12 analysis software was used to manage themes and codes.

Author Reflexivity

Prior to data collection, the interviewer and research team had minimal contact with participants (stakeholder engagement session). The participants were informed that the purpose of the study was to gather their views and experiences on e-Tracker use to assess how the immunization e-Tracker responds to stakeholders' needs and identifies areas for improvement. They

XSL•FO
RenderX

were also informed that this was part of a larger project studying the design and implementation of DHIs to improve childhood immunization in Rwanda. Authors entered this study with the belief that an e-Tracker had the potential to positively impact care providers' experiences; however, it took effort to prevent personal bias during data analysis. Due to COVID-19 restrictions, member checking was completed with 1 key informant from each category.

Ethical Considerations

This study was approved by the Rwanda National Ethics Committee (1011/RNEC/2020), the Norwegian (West) Regional Committee for Medical and Health Research Ethics (251925), and the Rwanda Ministry of Health's National Health Research Committee (reference NHRC/2021/PROT/002). All methods were performed in accordance with relevant guidelines and regulations by the World Medical Association Declaration of Helsinki—ethical principles for medical research involving human subjects [23].

The participants were informed about the study objectives, their voluntary participation, and their right to refuse participation at any time. The written informed consent form was obtained from each participant after getting an explanation about the research purpose and confirming their participation in the study. The interviews took place in a safe room with the office door locked at the health facility. The recorded information was transcribed and anonymized. The audio recording device could only be accessed via a security code by the lead author (TU).

Results

Overview

In total, 14 e-Tracker users were interviewed (Table 3), including 5 immunization nurses, 5 data managers, and 4 EPI supervisors (1 EPI supervisor declined being interviewed due to clinical COVID-19 work). Most of the immunization nurses were female (4/5, 80%) and had more than 10 years of work experience (3/5, 60%). In contrast, data managers were mostly male (4/5, 80%), younger, and had work experience of 5 years or less (4/5, 80%). Half of the supervisors (2/4, 50%) were female. The supervisors had varying levels of work experience (Table 3). We present our findings based on the constructs of the HOT-Fit framework.

Table 3. Characteristics of study participants.

Characteristics	Immunization nurses (n=5)	Data managers (n=5)	Supervisors (n=4)
Sex, n (%)			
Male	1 (20)	4 (80)	2 (50)
Female	4 (80)	1 (20)	2 (50)
Age range (years), n (%)			
25-35	1 (20)	2 (40)	1 (25)
36-45	2 (40)	3 (60)	0 (0)
46-55	2 (40)	0 (0)	2 (50)
56 and older	0 (0)	0 (0)	1 (25)
Age (years), mean (SD)	43 (8.29)	35 (5.89)	50 (10.23)
Field of study, n (%)			
Nursing	5 (100)	2 (40)	0 (0)
Laboratory	0 (0)	1 (20)	0 (0)
Computer science	0 (0)	1 (20)	0 (0)
Public health	0 (0)	1 (20)	3 (75)
Midwifery	0 (0)	0 (0)	1 (25)
Working experience (years), n (%)			
≤5	1 (20)	4 (80)	2 (50)
6-10	1 (20)	1 (20)	1 (25)
>10	3 (60)	0 (0)	1 (25)

Technology

System Quality

Data managers and supervisors reported that the e-Tracker was not a complex system. Two (40%) of 5 nurses perceived the e-Tracker as complex due to limited skills of computer literacy (Table 4: section A, construct 1).

XSL•FO
RenderX

Table 4. Summary of main findings in accordance with the HOT-Fit framework and quotes from key informant interviews.

Construct number and main findings	User's quotes
Section A: Technology	
System quality	
1. Ease of learning	"...that system [e-Tracker] is not difficult to use, except that it is not easy for everyone because there are some health centers for example that have immunization nurses who do not know how to use the computer." (EPI[a] supervisor 1)
2. Better data security than paper registers and forms	"e-Tracker is a secure system protected by personal credentials; it is not like paper registers where anyone can access." (Data manager 5)
3. Missing technical functionalities	"...as a person who is in the field and using it [e-Tracker] frequently, I realize that there are some functionalities that the e-Tracker is lacking. For example, it does not show me the next appointment for someone's vaccination or the list of who the nurses should be seeing today." (Data manager 5)
4. Not compatible for community outreach	"...internet connection that is not available, lack of outreach support—all these are challenges with using the e-Tracker." (Immunization nurse 1)
5. Connectivity issues and slow system response	"Things related to e-Tracker are slow, definitely slow. This is a challenge we usually face." (Immunization nurse 2)
Information quality	
6. Incomplete and unreliable data	"There are times when you register a child and when you go back to search him or her, you find that the actual information is not complete, or you find that the e-Tracker contains a duplicate of the child's records." (Immunization nurse 3)
7. Increased documentation workload	"I may fail to get time for instance, and they shift me to provide another health service, but, because there is much information that needs to be entered and I am responsible for that, I go quickly and take like one hour after work, or I come early in the morning to enter them." (Immunization nurse 4)
Service quality	
8. Delays in getting technical support	"It is difficult to get technical assistance because it is from central level and nowhere else...if the problem is simple like the system is off and then back on, those ones are quick and can be done on a phone call or WhatsApp, but bigger technical issues take time." (Data manager 2)
9. Alternative communication lines	"...talking about the other [communication] chain...I just call my superior at the hospital, and he conveys it to the central level technical team...and they gradually communicate with each other, and the information reaches us." (Data manager 4)
Section B: Human	
System use	
10. Does not meet the intended purpose	"What I expected from e-Tracker up to now, I can say that I have not yet seen its results. This may be due to other challenges, but the functionalities required by the nurses to use the e-Tracker well and properly are not yet available." (Supervisor 1)
11. Suboptimal use due to increased documentation workload	"This e-Tracker system is expected to be used by immunization nurses; it has apparently increased their work, which was not easy. That is simply to say, this is beyond their capacity." (Supervisor 1)
User satisfaction	
12. General dissatisfaction with the e-Tracker	"...Discriminating children's cards increases job, in e-Tracker it is simple; just search child and find him easily, but the use of e-Tracker did not stop papers, you complete all existing paper books and forms and then go complete e-Tracker." (Immunization nurse 2)
Section C: Organization	
Structure	
13. Lack of effective training processes	"...data manager who received training has gone, the one who replaced him does not actually know to use e-Tracker, he often called me asking, 'where can I click on?'...you realize that it is slowly by slowly." (Immunization nurse 2)
14. Lack of support for health workers in using technology	"...it happens that you register a child and when you go back to search for him [in the e-Tracker], you miss him simply because you do not know if it is a connection problem, or a low knowledge regarding how to search for him." (Immunization nurse 1)
Environment	
15. Performance-based financing	"We have many duties, and there are so many systems at health center...they come and say we give you PBF after seeing in the system how many children you have entered, and it is understandable that you will not receive any money if you didn't register any child." (Immunization nurse 3)

XSL•FO
RenderX

Construct number and main findings	User's quotes
Section D: Net benefits	
16. Perceived current benefits	"...the e-Tracker has a dashboard for data analysis. Like now, I sometimes say, let me see how many children we have registered this month; for the first, the second and the third dose of Penta, for instance." (Data manager 3)
17. Perceived future benefits for HWs[b]	"e-Tracker has made nothing easier for me. Instead, it has complicated things. Perhaps there is value in the e-Tracker if all these papers and books are removed. Then you may find that the e-Tracker will bring benefits." (Immunization nurse 3)

[a]EPI: Expanded Program on Immunization.

[b]HW: health worker.

Data security in the e-Tracker was generally perceived as satisfactory and better than data security using paper registers (Table 4: section A, construct 2). However, users reported several shortcomings. They cited the lack of several technical functionalities such as client lists, lists of defaulters, unspecific search and register functions, automated routine reports, and linkage to other systems such as birth notification and the national identity system (Table 4: section A, construct 3). Users expressed the need for a more flexible data entry tool that can operate offline, such as handheld tablets instead of desktop computers, to use during community outreach. They also cited poor connectivity and solely relying on health center–purchased internet as one of the most important reasons for the suboptimal use of the e-Tracker (Table 4: section A, construct 4). For system response time, 4 (80%) of 5 health workers and 4 (80%) of 5 data managers reported that the e-Tracker responds slowly. The remaining interviewees, particularly supervisors, located at hospitals with better internet connectivity, reported the opposite that the e-Tracker had a quick response time. Adequate support for network connectivity was lacking (Table 4: section A, construct 5). For example, immunization health workers at health centers were given modems, but they claimed that they were not given financial assistance for continued internet subscriptions.

Information Quality

Data in the e-Tracker were considered incomplete and unreliable and were not actively used by the immunization nurses (Table 4: section A, construct 6). Several underlying issues were identified as contributors to poor information quality. Users were required to document in the e-Tracker in addition to existing paper forms, which created double work. The double entry of data, combined with a mismatch between the data elements in the paper forms and the e-Tracker, results in users skipping some data fields in the e-Tracker. A common response with all users was the lack of time to complete documentations in the e-Tracker due to heavy workloads (Table 4: section A, construct 7). Two (40%) of the 5 interviewed immunization nurses were not trained in e-Tracker use, but even those who were trained and able to use the e-Tracker reported that the time allocated to them to fill the e-Tracker was insufficient. Three (60%) of 5 immunization nurses reported having to work overtime to enter data in the e-Tracker, 1 hour before or after work—a practice that users believed adversely affected data quality.

Service Quality

All interviewed nurses and data managers reported some form of delay in getting technical support (Table 4: section A, construct 8). Users' responses on this issue suggest that they might prefer reporting issues to their supervisor, who could then facilitate communication with the central support team.

Human

System Use

According to all interviewees, the e-Tracker did not meet overall user expectations. Further exploration revealed that users want a system that generates automated monthly reports and reduces documentation workload. The e-Tracker does not automatically generate any reports, and double documentation was identified as an important problem that impacted effective e-Tracker use (Table 4: section B, constructs 10, 11).

User Satisfaction

When asked whether they were satisfied with the e-Tracker, only 2 (40%) of 5 data managers said yes. The lack of technical functionalities and increased documentation workload were the leading causes of dissatisfaction for the data managers and health workers, respectively (Table 4 section B, construct 12).

Organization

Structure

Users described quarterly data quality assessment workshops to encourage e-Tracker use by health workers. Such assessments are usually done by data managers, nurses, and their supervisors by reviewing paper reports and e-Tracker reports and comparing them to HMIS reports for selected vaccination indicators, such as BCG. Health workers reported not receiving enough support in navigating digital systems in general (Table 4 section C, construct 13) and highlighted the need for regular training sessions on how to use the e-Tracker and a plan to deal with staff turnover. The planned training for users in 2021 did not happen due to the COVID-19 pandemic.

Environment

Performance-based financing was provided to the health workers based on the number of newborn babies registered as BCG vaccinated in comparison with their reported number of BCG vaccinations. The interviewees alluded to this as a reason for entering data for this specific indicator into the e-Tracker rather than the indicators for other vaccines. Performance-based financing in this context is based on the number of children

XSL•FO
RenderX

registered with BCG vaccination as a way of promoting the registration of newborn babies in the childhood immunization e-Tracker [24] (Table 4: section C, construct 15).

Net Benefits

Participants acknowledged the potential benefits of an e-Tracker provided technical and implementation issues are addressed. For example, all EPI supervisors reported that the tool could be helpful to monitor children's registration and vaccination status without visiting health centers physically. Two (40%) of 5 data managers reported using the e-Tracker for monitoring and evaluation in terms of vaccination coverage for their respective health centers. In contrast, all health workers did not report any net benefits from the current use, although they see that the e-Tracker may contribute positively to their work in the future (Table 4: section D, construct 17).

Key Improvements

Textbox 1 provides a summary of the main recommendations for improvement of the e-Tracker based on our findings.

Textbox 1. Overall recommendations for key improvements highlighted by the users.

Immunization nurses

- Better client search and register function
- Produce lists of expected and missed clients to avoid searching in paper registers
- Facilitate tracking a defaulter or a dropout child and remind parents of the missed appointments
- Improve connectivity
- Offline e-Tracker version that will make it easier to collect data in case of network outage, handheld devices to help immunization outreach in difficult-to-reach areas
- Regular training on e-Tracker use

Data managers

- Generate automatic monthly reports
- Link e-Tracker to other systems such as Civil Registration and Vital Statistics and national identification systems

Expanded Program on Immunization supervisors

- Additional trainings on analysis of e-Tracker data
- Offline e-Tracker version and more devices to support nurses' work at primary health centers

Discussion

Principal Findings

This study explored stakeholders' experiences and perceptions of using the e-Tracker for the Rwandan childhood immunization program. Users of the e-Tracker described several issues that hamper effective data entry as well as data use. Data in e-Tracker were reported to be incomplete and unreliable as result of dual documentation on paper and digitally.

Rwanda is one of the few countries in Africa to implement an EIR at scale. Implementation of the e-Tracker is a top priority for the childhood immunization program. Along with technological resources such as computers and modems, a top-level team and 3 cadres of trained health professionals from each health center across the nation are assigned to support the implementation indicating significant organizational support for change. EIRs allow for real-time monitoring of immunization status and provide data for decision-making, and their evaluations play a key role in identifying strategies to improve their use [25]. Our findings demonstrate the need for technical improvements to fit clinical practice and increase benefits, addressing implementation-related issues such as workflow matching, as well as training and user support. User-informed development of technical functionalities has been shown to be linked to higher adoption of health information systems in a systematic review of 55 studies [26]. Slow response times and delayed IT support adversely affected e-Tracker use in our study, factors also reported in other studies of digital information systems [27,28].

Creating an enabling environment for digital health systems by addressing issues such as training, and capacity strengthening in data entry and use, is equally important to ensure successful implementation [29]. Users cited a general dissatisfaction with the e-Tracker for several reasons including increased workload due to dual documentation and insufficient training. Several studies have reported similar dissatisfaction among users of digital health information systems in many cases as a result of the system's inability to match existing work patterns [26]. On the other hand, users are typically more satisfied when information systems offer good quality data; the higher the quality of the data the higher the satisfaction [27,30]. Users in this study perceived the information in the e-Tracker to be inaccurate and incomplete in comparison with the paper records and registers. None of the entered digital information was used by data managers or nurses for clinical practice.

Immunization nurses are the intended users of the e-Tracker, although the current workflow involves secondary data entry in the e-Tracker by the data manager in several health centers. While data managers and supervisors stated some benefits of the e-Tracker for their work, immunization nurses reported no

XSL•FO
RenderX

net benefits of the e-Tracker as it has been implemented in its current version. One of the reasons for this may be that the e-Tracker in its current form is not considered an essential part of the data ecosystem in the immunization health information system, particularly because the monthly reports are still paper based and not generated from the e-Tracker. In a setting such as Rwanda with scarce human resources for health, efficiency and costs are important considerations. Efficiency gains cannot be achieved unless health centers phase out paper immunization records and exclusively use the e-Tracker for data entry [31]. Similarly, a study conducted in Zambia and Tanzania showed that the use of the EIR decreased over time in settings where it was used in parallel with paper-based documentation compared to exclusive use [32]. In most other LMICs, paper-based documentation and reporting consume a significant proportion of health workers' time, which can be alleviated by well-implemented digital tools co-designed with the end user [31,33].

Organizations play a key role in supporting the adoption of digital systems directly and indirectly and sometimes inadvertently skewing priorities [20,30]. For instance, in our study, health workers are provided with performance-based financing based on BCG vaccine coverage rates, which might explain the relatively better completeness of these data in the e-Tracker.

Strengths

This study was conducted in sub-Saharan Africa, where there has been relatively limited research on EIRs and DHIs in general. Our findings are reasonably generalizable to the Rwandan context for two main reasons: (1) we sampled health centers at different stages of e-Tracker use, ranging from low to high, and (2) we included all users of the e-Tracker (immunization nurses, data managers, and supervisors).

Most studies that have applied the HOT-Fit framework have used quantitative methods to evaluate the effectiveness. We chose qualitative methods to gain an in-depth understanding of user-reported barriers and opportunities for e-Tracker use [20]. Our research is aligned with the national health system priorities to improve data use in the immunization program [34]. Key stakeholders, including representatives from the Ministry of Health and the Rwanda Biomedical Center, were involved at every stage of the research. They were consulted and presented with the study plan and results.

Study Limitations

This study has some limitations. The study was conducted in 2021, after a relatively short period of e-Tracker use by the health centers. Since the first introduction of the tool, some improvements have been implemented and these were not captured in our study. For example, nationwide linkages between the CRVS and immunization registry have recently been established and health workers providing immunization can retrieve information about the child from the CRVS. The COVID-19 pandemic and the subsequent restrictions in the years following the implementation of the e-Tracker may have affected training, use, and perceptions. Health workers from the immunization program (immunization nurses, data managers, and EPI supervisors) contributed immensely to the COVID-19 response, which may have affected their attitudes and perceptions toward their general workload and e-Tracker use.

Conclusions

The study findings revealed a low satisfaction level among the users of the immunization e-Tracker in Rwanda due to technical as well as implementation-related factors. Technical functionalities and implementation strategies co-designed with the user can help improve user experience and eventually maximize the benefits of the e-Tracker. Implementation strategies to reduce or remove dual documentation on paper and digital systems and to generate automated digital monthly immunization reports can save valuable time for health workers.

Acknowledgments

The authors thank the Ministry of Health and Rwanda Biomedical Center for their continued collaboration and facilitation of the research. They are also grateful to the participating health centers and the cooperation of study participants. They also acknowledge Michael Mugisha for his invaluable contribution to the conception and design of this research effort.

Conflicts of Interest

None declared.

Multimedia Appendix 1
COREQ (Consolidated Criteria for Reporting Qualitative Research) checklist.
[PDF File (Adobe PDF File), 536 KB - humanfactors_v11i1e53071_app1.pdf]

Multimedia Appendix 2
Interview guides.
[PDF File (Adobe PDF File), 125 KB - humanfactors_v11i1e53071_app2.pdf]

References

XSL•FO
RenderX

1. Global immunization coverage 2021. World Health Organization. 2021. URL: https://www.who.int/en/news-room/fact-sheets/detail/immunization-coverage [accessed 2022-10-26]
2. Carter A, Msemburi W, Sim SY, Gaythorpe KAM, Lambach P, Lindstrand A, et al. Modeling the impact of vaccination for the immunization agenda 2030: deaths averted due to vaccination against 14 pathogens in 194 countries from 2021 to 2030. Vaccine 2023;S0264-S0410X:00854-X [FREE Full text] [doi: 10.1016/j.vaccine.2023.07.033] [Medline: 37537094]
3. Tozzi AE, Gesualdo F, D'Ambrosio A, Pandolfi E, Agricola E, Lopalco P. Can digital tools be used for improving immunization programs? Front Public Health 2016;4:36 [FREE Full text] [doi: 10.3389/fpubh.2016.00036] [Medline: 27014673]
4. Murray E, Hekler EB, Andersson G, Collins LM, Doherty A, Hollis C, et al. Evaluating digital health interventions: key questions and approaches. Am J Prev Med 2016;51(5):843-851 [FREE Full text] [doi: 10.1016/j.amepre.2016.06.008] [Medline: 27745684]
5. World Health Organization. WHO Guideline. Recommendations on Digital Interventions for Health System Strengthening. Geneva: World Health Organization; 2019. URL: https://iris.who.int/bitstream/handle/10665/311941/9789241550505-eng.pdf?sequence=31
6. Closing gaps in vaccine coverage with electronic immunization registries. PATH. URL: https://www.path.org/case-studies/electronic-immunization-registries/ [accessed 2023-01-09]
7. Namageyo-Funa A, Samuel A, Bloland P, Macneil A. Considerations for the development and implementation of electronic immunization registries in Africa. Pan Afr Med J 2018;30:81 [FREE Full text] [doi: 10.11604/pamj.2018.30.81.11951] [Medline: 30344865]
8. Peters DH, Tran NT, Adam T, Alliance for Health Policy and Systems Research. Implementation research in health: a practical guide. World Health Organization. 2013. URL: https://ahpsr.who.int/publications/i/item/9789241506212 [accessed 2024-04-25]
9. Parajuli R, Bohara D, Kc M, Shanmuganathan S, Mistry SK, Yadav UN. Challenges and opportunities for implementing digital health interventions in Nepal: a rapid review. Front Digit Health 2022;4:861019 [FREE Full text] [doi: 10.3389/fdgth.2022.861019] [Medline: 36120714]
10. Comprehensive multi-year plan 2013-2017. Rwanda Ministry of Health. 2012. URL: http://www.nationalplanningcycles.org/sites/default/files/country_docs/Rwanda/attachment_6_revised_cmyp_08.pdf [accessed 2024-04-25]
11. Joint appraisal report 2017 Rwanda. GAVI. 2017. URL: https://www.gavi.org/country/rwanda/documents/ [accessed 2024-04-25]
12. Maïga A, Jiwani SS, Mutua MK, Porth TA, Taylor CM, Asiki G, et al. Generating statistics from health facility data: the state of routine health information systems in Eastern and Southern Africa. BMJ Glob Health 2019;4(5):e001849. [doi: 10.1136/bmjgh-2019-001849] [Medline: 31637032]
13. Nwankwo CM, Orua E. Factors influencing incomplete immunization among under five years old children at CHUK Hospital, Nyarugenge district, Rwanda. Int J Community Med Public Health 2020;7(10):3787-3795 [FREE Full text] [doi: 10.18203/2394-6040.ijcmph20204340]
14. RWANDA - EPI Tracker and Dashboard use case. URL: https://s3-eu-west-1.amazonaws.com/content.dhis2.org/general/EPI+Tracker+%26+Dashboard+Implementation_+Rwanda.pdf [accessed 2024-04-25]
15. Republic of Rwanda, Ministry of Health—Fourth Health Sector Strategic Plan July 2018-June 2024. Rwanda Ministry of Health. 2018. URL: http://npngti.com/wp-content/uploads/2018/06/Rwanda_Nat-Health-Sector-Plan_2018-2024.pdf [accessed 2024-04-25]
16. University of Oslo. The world's largest health information management system — developed through global collaboration led by UiO. Dhis2. URL: https://dhis2.org/ [accessed 2023-02-12]
17. Hamilton AB, Finley EP. Qualitative methods in implementation research: an introduction. Psychiatry Res 2019;280:112516 [FREE Full text] [doi: 10.1016/j.psychres.2019.112516] [Medline: 31437661]
18. Gu Y, Warren J. Chapter 11 Methods for descriptive studies. In: Lau F, Kuziemsky C, editors. Handbook of eHealth Evaluation: An Evidence-based Approach. Victoria, British Columbia Canada V8P 5C2: University of Victoria; 2016.
19. Stetler CB, Legro MW, Wallace CM, Bowman C, Guihan M, Hagedorn H, et al. The role of formative evaluation in implementation research and the QUERI experience. J Gen Intern Med 2006;21(Suppl 2):S1-S8 [FREE Full text] [doi: 10.1111/j.1525-1497.2006.00355.x] [Medline: 16637954]
20. Yusof MM, Kuljis J, Papazafeiropoulou A, Stergioulas LK. An evaluation framework for health information systems: Human, Organization and Technology-fit factors (HOT-fit). Int J Med Inform 2008;77(6):386-398. [doi: 10.1016/j.ijmedinf.2007.08.011] [Medline: 17964851]
21. Tong A, Sainsbury P, Craig J. Consolidated criteria for reporting qualitative research (COREQ): a 32-item checklist for interviews and focus groups. Int J Qual Health Care 2007;19(6):349-357 [FREE Full text] [doi: 10.1093/intqhc/mzm042] [Medline: 17872937]
22. Meunier PY, Raynaud C, Guimaraes E, Gueyffier F, Letrilliart L. Barriers and facilitators to the use of clinical decision support systems in primary care: a mixed-methods systematic review. Ann Fam Med 2023;21(1):57-69 [FREE Full text] [doi: 10.1370/afm.2908] [Medline: 36690490]

XSL•FO
RenderX

23. World Medical Association. World Medical Association declaration of Helsinki. Ethical principles for medical research involving human subjects. Bull World Health Organ 2001;79(4):373-374 [FREE Full text] [Medline: 11357217]
24. Performance based financing procedures manual for health facilities (Hospitals and Health Centers). Rwanda Ministry of Health. 2021. URL: https://www.scribd.com/document/529709575/1562825127078-PBF-Manual-Hospital-and-Health-Center [accessed 2024-04-25]
25. Deharja A, Hargono A, Santi MW, Nandini N, Damayanti NA. Evaluating the usability of Hospital Information System (HIS) through Human Organization Technology-Fit (Hot-Fit) model. Int Proc 2nd Int Sci Meet Health Inf Manag 2020;5(1171):380-389 [FREE Full text]
26. Yusof MM, Stergioulas L, Zugic J. Health information systems adoption: findings from a systematic review. Stud Health Technol Inform 2007;129(Pt 1):262-266. [Medline: 17911719]
27. Fauzan A, Noviandi N. Evaluation of optima regional health information system with HOT-Fit on technology aspects approach in Johar Baru Health Center Jakarta. J Intell Comput Health Informatics 2020;1(1):9-14 [FREE Full text] [doi: 10.26714/jichi.v1i1.5397]
28. Erlirianto LM, Ali AHN, Herdiyanti A. The implementation of the Human, Organization, and Technology–Fit (HOT–Fit) framework to evaluate the Electronic Medical Record (EMR) system in a hospital. Procedia Comput Sci 2015;72:580-587 [FREE Full text] [doi: 10.1016/j.procs.2015.12.166]
29. Pan American Health Organization. Electronic immunization registry: practical considerations for planning, development, implementation and evaluation. Pan American Health Organization. 2017. URL: https://iris.paho.org/handle/10665.2/34865 [accessed 2024-04-25]
30. Kilsdonk E, Peute LW, Jaspers MWM. Factors influencing implementation success of guideline-based clinical decision support systems: a systematic review and gaps analysis. Int J Med Inform 2017;98:56-64. [doi: 10.1016/j.ijmedinf.2016.12.001] [Medline: 28034413]
31. Venkateswaran M, Nazzal Z, Ghanem B, Khraiwesh R, Abbas E, Khader KA, et al. eRegTime-Time spent on health information management in primary health care clinics using a digital health registry versus paper-based documentation: cluster-randomized controlled trial. JMIR Form Res 2022;6(5):e34021 [FREE Full text] [doi: 10.2196/34021] [Medline: 35559792]
32. Carnahan E, Ferriss E, Beylerian E, Mwansa FD, Bulula N, Lyimo D, et al. Determinants of facility-level use of electronic immunization registries in Tanzania and Zambia: an observational analysis. Glob Health Sci Pract 2020;8(3):488-504 [FREE Full text] [doi: 10.9745/GHSP-D-20-00134] [Medline: 33008860]
33. World Health Organization (WHO). Web Annex M: Cross-cutting acceptability and feasibility issues and implementation considerations for health workers. In: WHO guideline Recommendations on Digital Interventions for Health System Strengthening. Geneva: World Health Organization; 2019:1-5.
34. Mutale W, Chintu N, Amoroso C, Awoonor-Williams K, Phillips J, Baynes C, et al. Improving health information systems for decision making across five sub-Saharan African countries: implementation strategies from the African Health Initiative. BMC Health Serv Res 2013;13(Suppl 2):S9 [FREE Full text] [doi: 10.1186/1472-6963-13-S2-S9] [Medline: 23819699]

Abbreviations

BCG: Bacille Calmette-Guérin
COREQ: Consolidated Criteria for Reporting Qualitative Research
CRVS: Civil Registration and Vital Statistics
DHI: digital health intervention
EIR: electronic immunization registry
EPI: Expanded Program on Immunization
HMIS: health management information system
HOT-Fit: Human, Organization, and Technology–Fit
LMIC: low- and middle-income country

Edited by A Kushniruk; submitted 25.09.23; peer-reviewed by JR Medina, D Chrimes, S Fenton; comments to author 24.01.24; revised version received 20.03.24; accepted 07.04.24; published 28.05.24.

Please cite as:
Uwera T, Venkateswaran M, Bhutada K, Papadopoulou E, Rukundo E, K Tumusiime D, Frøen JF
Electronic Immunization Registry in Rwanda: Qualitative Study of Health Worker Experiences
JMIR Hum Factors 2024;11:e53071
URL: https://humanfactors.jmir.org/2024/1/e53071
doi:10.2196/53071
PMID:38805254

XSL•FO
RenderX

©Thaoussi Uwera, Mahima Venkateswaran, Kiran Bhutada, Eleni Papadopoulou, Enock Rukundo, David K Tumusiime, J Frederik Frøen. Originally published in JMIR Human Factors (https://humanfactors.jmir.org), 28.05.2024. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on https://humanfactors.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

The Solutions in Health Analytics for Rural Equity Across the Northwest (SHARE-NW) Dashboard for Health Equity in Rural Public Health: Usability Evaluation

Elizabeth Heitkemper[1*], RN, PhD; Scott Hulse[2*], MD, MPH, MSc, MSc; Betty Bekemeier[3,4], MPH, PhD; Melinda Schultz[3], MSc; Greg Whitman[3], BA; Anne M Turner[4,5], MD, MPH, MLIS

[1]School of Nursing, The University of Texas at Austin, Austin, TX, United States
[2]School of Medicine, Uniformed Services University of the Health Sciences, Bethesda, MD, United States
[3]School of Nursing, University of Washington, Seattle, WA, United States
[4]School of Public Health, University of Washington, Seattle, WA, United States
[5]School of Medicine, University of Washington, Seattle, WA, United States
[*]these authors contributed equally

Corresponding Author:
Elizabeth Heitkemper, RN, PhD
School of Nursing
The University of Texas at Austin
1710 Red River Street
Austin, TX, 78712
United States
Phone: 1 512 232 4228
Email: e.heit@utexas.edu

Abstract

Background: Given the dearth of resources to support rural public health practice, the solutions in health analytics for rural equity across the northwest dashboard (SHAREdash) was created to support rural county public health departments in northwestern United States with accessible and relevant data to identify and address health disparities in their jurisdictions. To ensure the development of useful dashboards, assessment of usability should occur at multiple stages throughout the system development life cycle. SHAREdash was refined via user-centered design methods, and upon completion, it is critical to evaluate the usability of SHAREdash.

Objective: This study aims to evaluate the usability of SHAREdash based on the system development lifecycle stage 3 evaluation goals of efficiency, satisfaction, and validity.

Methods: Public health professionals from rural health departments from Washington, Idaho, Oregon, and Alaska were enrolled in the usability study from January to April 2022. The web-based evaluation consisted of 2 think-aloud tasks and a semistructured qualitative interview. Think-aloud tasks assessed efficiency and effectiveness, and the interview investigated satisfaction and overall usability. Verbatim transcripts from the tasks and interviews were analyzed using directed content analysis.

Results: Of the 9 participants, all were female and most worked at a local health department (7/9, 78%). A mean of 10.1 (SD 1.4) clicks for task 1 (could be completed in 7 clicks) and 11.4 (SD 2.0) clicks for task 2 (could be completed in 9 clicks) were recorded. For both tasks, most participants required no prompting—89% (n=8) participants for task 1 and 67% (n=6) participants for task 2, respectively. For effectiveness, all participants were able to complete each task accurately and comprehensively. Overall, the participants were highly satisfied with the dashboard with everyone remarking on the utility of using it to support their work, particularly to compare their jurisdiction to others. Finally, half of the participants stated that the ability to share the graphs from the dashboard would be “extremely useful” for their work. The only aspect of the dashboard cited as problematic is the amount of missing data that was present, which was a constraint of the data available about rural jurisdictions.

Conclusions: Think-aloud tasks showed that the SHAREdash allows users to complete tasks efficiently. Overall, participants reported being very satisfied with the dashboard and provided multiple ways they planned to use it to support their work. The main usability issue identified was the lack of available data indicating the importance of addressing the ongoing issues of missing and fragmented public health data, particularly for rural communities.

XSL•FO
RenderX

(JMIR Hum Factors 2024;11:e51666) doi:10.2196/51666

KEYWORDS

data dashboard; rural health; health equity; usability; nursing informatics; dashboard; rural; informatics; satisfaction; think aloud; content analysis; user experience; public health; visualization; information systems

Introduction

Data visualization dashboards developed to address health and equity have become increasingly popular [1,2]. Leveraging the longstanding history of using dashboards to aggregate and analyze data in public health [3] and medicine [4], these new dashboards cover myriad health equity–focused topics and target broad audiences. Recently, Thorpe and Gourevitch [5] identified 15 examples of US-based health dashboards that illustrate this growing trend. Examples range from a COVID-19 dashboard that highlights inequities in cases and deaths by geography to a policy dashboard that aggregates local laws and policies that affect population health [5]. Similar to these dashboards, the solutions in health analytics for rural equity across the northwest (SHARE-NW) dashboard (SHAREdash) was created to address health equity for rural communities.

Delivery and allocation of health services through public health agencies is a key mechanism for achieving health equity in the United States as they provide health prevention and promotion services and care [6]. Nationally, people in rural and frontier jurisdictions have significant health disparities compared with urban populations but are frequently the least well served by their public health agencies—local health departments (LHDs) [7,8]. Exacerbating this is the poor public health data systems, as updating to include information on structural and social factors has not been a top priority in LHDs' activities or spending [8,9]. Research has highlighted the critical need to improve timely and reliable population health data to inform resource allocation and decision-making [10-14]. Consequently, decisions regarding the delivery of public health services and care primarily rely on conventional wisdom. This results in services that frequently do not reflect the needs of the populations they serve resulting in wasteful, harmful, and inequitable inefficiencies that exacerbate existing disparities [15-17]. To address these issues and support LHDs serving rural areas, the goal of SHAREdash is to provide accessible and relevant data that will enable public health professionals to identify, communicate, and address health disparities in their jurisdictions and with their communities. Developed with user-centered participatory design methods and guided by Munzer's Nested Model for Visualization Design and Validation [18], SHAREdash is the first rigorously designed health equity dashboard developed for rural communities that we are aware of [19].

While clear objectives and thoughtful design are critical to ensuring the development of useful dashboards, Thorpe and Gourevitch [5] highlight the importance of evaluating dashboards and the need for a more rigorous assessment of the effectiveness and usefulness of health equity dashboards. Evaluating the performance of a dashboard through end user usability testing is a critical and often missed component of dashboard creation. The International Organization for Standardization defines usability as "the extent to which a product can be used by specified users to achieve specified goals with effectiveness, efficiency, and satisfaction in a specific context of use" [20,21]. Poor usability has been shown to increase errors [22-24], increase the time to complete tasks [25], and reduce user uptake [26-28] and implementation efforts [29].

Proper assessment of technology usability should occur at multiple stages throughout the system development lifecycle (SDLC) and use the methods most appropriate for that respective stage. In a review of usability study methodologies of health information technology by Yen and Bakken [30], the authors outline the importance of conducting multiple usability evaluations that align with the 5 stages of the SDLC. Furthermore, the Yen and Bakken [30] review clarifies the differences in usability evaluation types and goals based on the SDLC stage of the technology (Multimedia Appendix 1). Results from SHAREdash's usability testing for SDLC stages 1 and 2 have been previously published [10,19]. Stage 2 findings were used to make critical changes and inform dashboard completion. Now that SHAREdash is finished and has entered SDLC stage 3, we evaluated its usability by examining all components combined (ie, the finished dashboard). Thus, the aim of this study was to evaluate the SDLC stage 3 evaluation goals of efficiency, satisfaction, and validity for SHAREdash.

Methods

The SHARE-NW Project and Dashboard

SHARE-NW is a partnership research project that was created with the goal of making data available and accessible to rural LHD practitioners, while building their capacity for data use and data-driven decision-making [10]. Partnering with LHDs in Alaska, Idaho, Oregon, and Washington, 7 priority topic areas (obesity, diabetes, tobacco use, mental and behavioral health, violence and injury prevention, oral health, and demographics) were identified during stage 1 of the SDLC for SHAREdash [19]. Data for the dashboard come from 36 unique data sources, including national data from the Centers for Disease Control and Prevention as well as local agencies and health departments (Multimedia Appendix 2). Data were deemed relevant to be included in the dashboard if it (1) addressed 1 of the 7 priority areas and (2) was provided at the county level so that it would be relevant to LHDs. To ensure the dashboard is usable and relevant for users, its features (eg, dynamic filters, pop-up tooltips, and visualizations) were created in collaboration with the staff from partner LHDs during SDLC stage 2. SHARE-NW has also developed a curated repository of web-based trainings and webinars, including new training modules developed in 2021 and launched in 2022 when gaps were found in the related training desired by practitioners. The new training modules developed use problem-based learning to teach audiences how to use and communicate data to promote health equity.

XSL•FO
RenderX

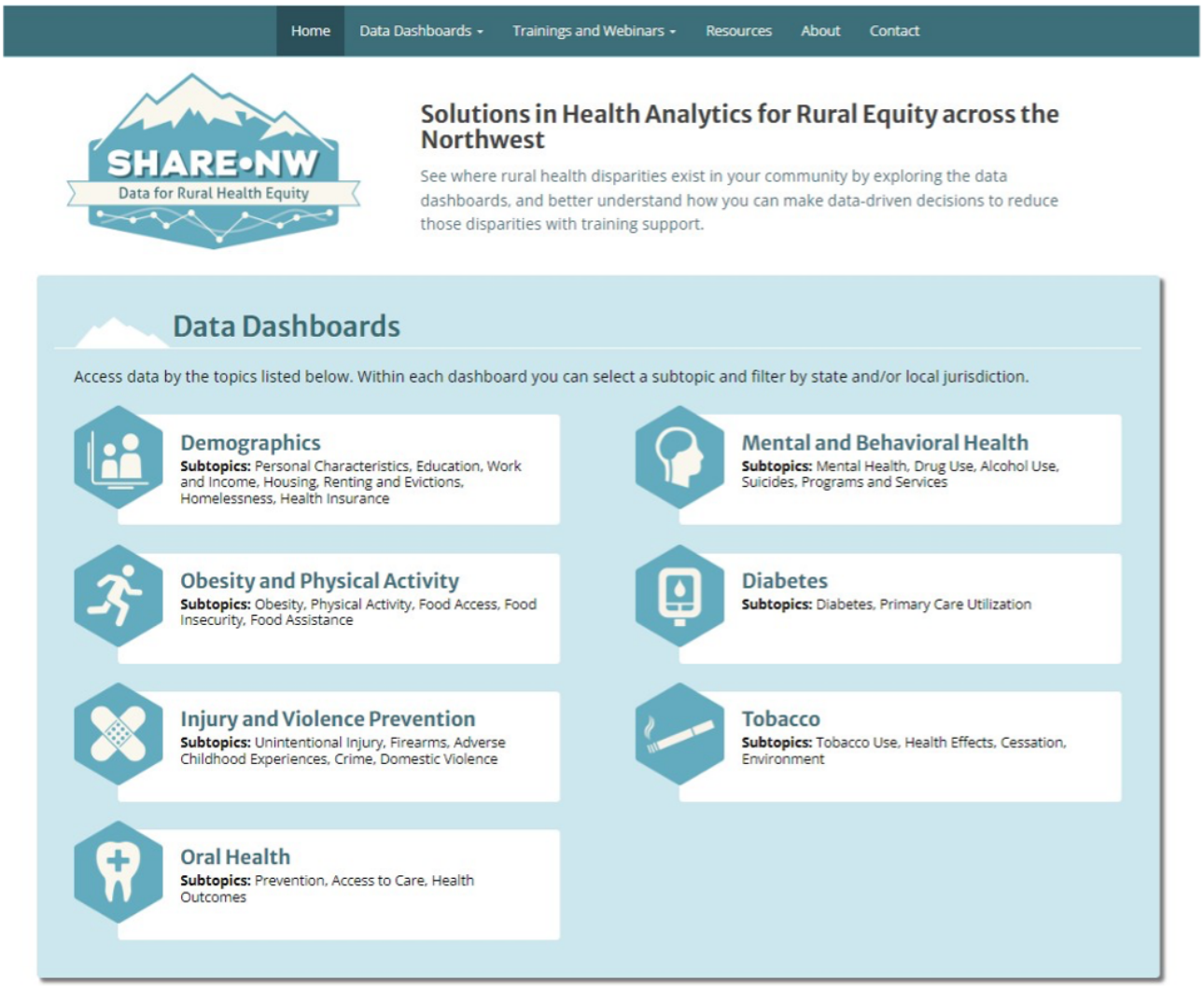
After conducting a needs assessment with rural LHD professionals during SDLC stage 1 [10], members of the SHARE-NW team identified a set of initial design requirements for SHAREdash. These requirements guided design and development decisions that ranged from key decisions, such as the selection of the best software to create the dashboard, to smaller decisions such as which size font to use for a graph label. Together with the findings from the SDLC stage 2 usability study, SHAREdash was completed and launched in August 2021.

SHAREdash Website and Interface

SHAREdash is a Tableau-based dashboard with a header at the top of the main page for users to locate information about the project and team, access resources on relevant topics such as data, communication, and health equity, how to contact a member of the team, and find the dashboards organized by priority topic area. Users can also see relevant trainings and webinars (both via drop-down boxes) on SHAREdash's main page. When users scroll down the main page, they can also find information on the website's purpose and design and see the sources of data powering the website. The largest feature on the main page (Figure 1) links to the 7 dashboards on the priority topic areas mentioned previously [19]. Within each dashboard, users can find state and county-level data organized by relevant subtopics. For example, the topic of "Tobacco" includes subtopics of "Tobacco use," "Health effects," "Cessation," and "Environment."

When users navigate to each of the main topics, they find a header that lists the main topic and each subtopic along the top, such that users can click through them. Within each subtopic, there are several drop-downs that allow users to filter the data. The primary drop-down prompts the user to "Select an Indicator." Some examples of indicators for the topic of "Tobacco" and the subtopic of "Tobacco use" are as follows: "8th graders' current tobacco use include Percent %," "high school students who smoked tobacco in the past month: Percent %," "Adults who currently smoke, Age-adjusted Percent %," and "Current e-cigarette use: Percent %." The remaining drop-downs allow users to filter the results by state, region, rurality (eg, rural or not rural), and jurisdiction types (eg, county). Users can export any of the dashboard views using the 3 options of export image, export to PDF, and share a link that is found along the top of the page. Along the bottom of each dashboard is the clickable link to view the data sources for this dashboard along with a statement explaining which data are listed as unavailable data within the dashboards.

Figure 1. The solutions in health analytics for rural equity across the northwest dashboard (SHAREdash) home page.



XSL•FO
RenderX

Study Setting and Participants

To evaluate this web-based dashboard, our study was conducted from January to April 2022. Participants were recruited from the states upon which SHAREdash was focused—Washington, Idaho, Alaska, and Oregon. All individuals who were rural public health professionals or trainees and had completed at least 1 prior SHARE-NW activity (2017-2022) and agreed to be contacted for future research activities (n=20). Prior SHARE-NW activities included the following: key informant interviews, interviews about the response to COVID-19, web-based surveys, dashboard mock-up testing sessions, dashboard usability testing activities, and group-based dashboard live training sessions. This ensured that all participants met the eligibility criteria of being at least 18 years old, working in public health for at least a year, and were in 1 of the 4 northwest states included in SHAREdash. Given public health differences by state, recruitment efforts ensured that at least 3 of the states were represented. Recruitment of this convenience sample had no additional inclusion or exclusion criteria.

Ethical Considerations

The University of Washington's institutional review board (IRB) approved the study protocol before participant recruitment (STUDY00013451). This study's IRB was approved as a modification to an original approval (MOD00011747; approved on December 13, 2021).

An initial recruitment email briefly summarizing the study purpose was sent to all prior participants in SHARE-NW activities. A follow-up recruitment email was sent 2 weeks after the initial email. The response rate for recruitment was 45% (n=9) with 1 person stating that they did not want to participate and the remaining 10 people not responding. Individuals who expressed interest were sent an email with information about the study, consent to participate, and instructions on scheduling their interview. Dashboard evaluation sessions consisted of 2 think-aloud tasks [31] and open-ended interview questions [32] regarding the participant's occupation and perceptions of the dashboard (Multimedia Appendices 3 and 4). Recruitment stopped when the following metrics were reached: (1) alignment with the published literature on the minimum number of participants needed to identify usability issues [33] and (2) when data saturation was reached. Given how stretched our public health partners were from the COVID-19 pandemic, the study team was cautious to not overburden them with study participation requests.

The think-aloud tasks served 2 purposes—the first was to refamiliarize the participant with SHAREdash and the second was to examine the usability components of effectiveness and efficiency. The first think-aloud task had the participant complete a simple task that consisted of switching between different subtopics, filtering for the participant's county, and changing the time frame being viewed. The second, more complex task included navigating to the right topic, filtering for a specific health outcome type, year, and rate, and identifying the original sources of the data being viewed. During testing, the moderator prompted participants to "think aloud" that is, verbalize their thoughts as they worked through the task. Following the tasks, the participants completed qualitative interviews that asked participants for their perceptions regarding SHAREdash's efficiency, validity, and satisfaction. The semistructured interview guide included questions that asked about the design aesthetics and functionality of SHAREdash, how quickly they are able to perform tasks, and the benefits and issues with using SHAREdash. The evaluation sessions were completed and recorded via a videoconferencing platform since screen sharing was needed for the 2 think-aloud task evaluations. Transcripts were automatically generated by the videoconferencing platform and stored securely in a password-protected cloud-based repository. A member of the research team deidentified and corrected any errors in the verbatim transcripts prior to analysis.

Recruitment and Data Collection

Think-Aloud Task Analysis

Operational definitions of the outcomes align with the International Organization for Standardization definitions of efficiency, satisfaction, and validity [34]. Efficiency and validity were primarily evaluated through the think-aloud tasks. The number of clicks taken to complete the task indicated efficiency and data on the participants' success of task completion were operationalized as "yes" (eg, no assistance needed), "no" (eg, assistance needed), or "partial" (eg, where the moderator confirms participant choices as either correct or incorrect but offers no other assistance) which indicated validity. Task 1 could be completed in a minimum of 7 clicks and task 2 could be done in 9 clicks. Transcripts were automatically generated and edited by a member of the research team for accuracy. Descriptive statistics were used to analyze the data and was completed in Excel. Quotes from the think-aloud tasks were analyzed to evaluate common efficiency issues, examine overall satisfaction, and assess validity. A control arm was not used in this study based on prior work that identified the inability of participants to complete these tasks without SHAREdash [10,19].

Qualitative Analysis

Data analysis of qualitative interview transcripts started with a directed, deductive approach to content analysis that was guided by a codebook comprising the initial codes of efficiency, satisfaction, and validity [35]. From this initial schema, iterative coding categories emerged as themes were developed. Coding was performed in NVivo (Lumivero) to organize data and provide an audit trail. Our interdisciplinary team of researchers met for an initial 90-minute collaborative coding session to talk through coding procedures and develop consensus for initial categories. Subsequent coding was performed independently with researchers meeting for 60-minute coding meetings to discuss categories and resolve discrepancies. Procedures for ensuring credibility, transferability, dependability, and confirmability were incorporated throughout the research process to ensure data trustworthiness. These procedures included taking field notes, team debriefing, reflexive journaling, consideration of negative cases, and maintenance of an audit trail. Data saturation was reached with researchers initially identifying the potential for saturation after the sixth participant interview and later confirming it with the ninth and final participant.

XSL•FO
RenderX

Results

Overview

Interviews lasting an average of 21 (SD 5.4) minutes were conducted between January and April 2022. Of the 9 public health practitioners interviewed, 4 were from Idaho, 3 were from Oregon, and 2 were from Washington (Table 1). Participants all identified as female, and the majority worked for health departments (n=8). Job positions included a director (n=2), managers (n=2), program specialists/coordinators (n=3), an epidemiologist (n=1), and a student/public health intern (n=1). Prior experience with Tableau was minimal with the majority (7/9, 78%) reporting less than 3 months experience to no experience.

Table 1. Demographics of dashboard evaluation participants (N=9).

Characteristic	Values, n (%)
State	
Idaho	4 (44)
Oregon	3 (33)
Washington	2 (22)
Organization type	
Local health department	7 (78)
State health department	1 (11)
Educational institution	1 (11)
Position	
Director	2 (22)
Coordinator	2 (22)
Manager	2 (22)
Other	3 (33)
Tableau experience	
0-3 months	7 (78)
>3 months	1 (11)
Not reported	1 (11)
Sex	
Female	9 (100)
Ethnicity	
Not reported	5 (56)
White	4 (44)

Think-Aloud Task Results

For efficiency, mean clicks were 10.1 (SD 1.4) for task 1 (with a minimum of 7 clicks) and 11.4 (SD 2.0) for task 2 (with a minimum of 9 clicks; Table 2). For task 1, extra clicks occurred when people tried to find the right place to filter for the correct dashboard page. One participant (participant 9) required partial assistance with one of the steps in the first task. They initially thought to navigate to the default subtopic of “Personal Characteristics” instead of the correct subtopic “Homelessness” in the “Demographics” dashboard. Although only 1 participant required assistance with this step, many participants took extra time with it. For task 2, extra clicks resulted from people looking for the sources of the data, which were located at the footer of each dashboard. Most participants were able to find the “View Data Sources” button easily because of the dashboard’s instructions or because it was where they expected it based on their prior experience. However, 3 participants noted that they naturally scrolled to the bottom of the dashboard looking for it but confused the “Resources” button with the “View Data Sources” button. This issue not only resulted in extra clicks but also was the point where these participants required confirmation by the moderator to continue. Of the 3 participants, 2 who had this issue stated that they were confused because they expected the data sources to be in a clickable pop-up or an in-text citation, rather than loading onto a separate page.

XSL•FO
RenderX

Table 2. Efficiency task analysis results (N=9).

Task type	Task 1	Task 2
Efficiency		
Clicks, mean (SD)	10.1 (1.4)	11.4 (2.0)
Clicks, median (IQR)	10.5 (9.0-11.0)	11.0 (9.5-12.8)
Validity, n (%)		
Successful	8 (89)	6 (67)
Partial	1 (11)	3 (33)
Not successful	0 (0)	0 (0)

Validity scores for both tasks were high with all participants (n=9) able to complete each task accurately and comprehensively. Most participants received a "yes" indicating that they did not require any prompting—89% (n=8) participants for task 1 and 67% (n=6) participants for task 2. For participants who did not receive a score of "yes," they only required the moderator to either confirm or deny their decisions prior to moving on, resulting in a score of "partial." None of the participants received a "no," indicating they could not finish the tasks.

Qualitative Results

The following 4 themes regarding efficiency were identified: "using the best terms and names to increase efficiency," "drop-down filters reduce efficiency," "minor navigation issues affect efficiency," and "learnability will increase efficiency over time". The primary issue that came up in the think-aloud tasks and the interviews was related to the dashboard labels and names that informed the theme of "using the best terms and names to increase efficiency." Multiple participants brought up the term "jurisdiction" and pointed out that it is less intuitive than the word "county." "I think 'jurisdictions' is obviously not wrong; it just would be a little bit more user-friendly to label it 'county'" (participant 1). Similarly, as identified in the second task analysis, 2 participants found the "Resources" button confusing and suggested renaming it to something more specific to mitigate this confusion.

For the second theme of "drop-down filters reduce efficiency," participants described how the functionality of the filters was difficult to navigate between some options due to the length of the drop-down boxes. For example, filtering to examine a single county requires users to search down a long drop-down list for the exact county they are looking for. Participants described this by saying, "Maybe it would be a nice feature to be able to type in a county versus the drop-down box or having to—well I guess you can 'select all' so you don't have to go through and click them all to select, but just those little things might make it easier" [participant 4]. Another participant described how they expected the interface to be like other software they are used to using such that it allows users to enter free text into a search bar and then, "…when you start typing things it only picks the things that match it" (participant 8).

For the third theme of "minor navigation issues affect efficiency," a few participants had difficulty locating the various subtopics within a dashboard, despite them being listed underneath each dashboard topic. For example, 1 participant looked for the "homelessness" indicator under the wrong subtopic.

> *I was initially thinking, 'Oh 'Homelessness' must be in one of these drop downs, because it was listed as a subtopic,' but then I glanced across the screen, and—you know—I saw 'Homelessness' up in this corner [with the other subtopics].* [participant 2]

Similarly, 3 participants eventually correctly identified that "Demographics" was the dashboard where they would find homelessness data, but they initially looked for the "homelessness" indicator under the "Housing" subtopic instead of the "Homelessness" subtopic. A participant suggested ways that the design of SHAREdash could be updated to more clearly indicate the subtopics.

> *I would think that [the indicator] is definitely going to be in Oral Health. It took a bit when I first looked at [SHAREdash] to realize that there were tabs (e.g., different subtopics). I think the size of the font and the fact that they are the same color as the bar makes it, so they are not standing out.* [participant 1]

These design suggestions were checked with some subsequent participants who agreed that changing the font size and color would help the subtopics stand out.

For the final theme of "learnability will increase efficiency over time," participants spoke about how quickly they were able to figure things out in SHAREdash and reported that with repeated use they thought they would quickly improve over time. Half of the participants stated that first-use learnability was high such that SHAREdash was easy to use the first time they tried. "I would say that there's not a lot of websites out there, where you can pick up on things that quickly. So, I immediately don't have any areas for improvement" [participant 2]. Whereas the remaining half of the participants stated that they felt like they would get progressively better at using SHAREdash over time.

The following 3 themes related to overall satisfaction were identified: "high potential to support work," "enables meaningful comparisons," and "needs more up-to-date data." For the theme of "high potential to support work," participants spoke positively about how much they liked SHAREdash and the myriad ways they could use SHAREdash's various features to support their work. One-third of the participants mentioned how unique and helpful it was to have the ability to export and share graphs.

XSL•FO
RenderX

> *I think it has a lot of features that aren't necessarily easily found [in other dashboards]…blowing it [SHAREdash] up to full screen, downloading it, sharing it—that's not necessarily common with dashboards, so I appreciate that…It could be really useful for like a grant application or demographic reporting for part of a program.* [participant 5]

Two participants mentioned that they might direct others to the dashboard so they could interact with data, and this was described as something that would be "extremely useful" and "super helpful" in their work. Finally, several participants identified specific types of work that SHAREdash would meaningfully support such as completing community health assessments:

> *We would definitely want to look at this in relation to the approach that we took with our CHA [community health assessment]. Most recently, I was trying to mine all of the data sources that are already in existence to inform it and see where some gaps were, and then we did primary data seeking based off those gaps instead of trying to reproduce data that's already in existence. And so, this [SHAREdash] would be a really great one-stop-shop to look at a lot of different ones at one time.* [participant 2]

> *I think it is already something that's on our radar when we talk about this CHA [community health assessment] that I mentioned. So, we're not here to duplicate efforts; let's use what's out there. And so, we'll probably refer to it [SHAREdash] for that.* [participant 5]

For the theme of "enables meaningful comparisons," all but 2 participants reported that they were highly satisfied with SHAREdash and cited the ability to compare their county or region with other neighboring or similar counties in different states as the reason why. Multiple participants stated they wanted to look at counties in nearby states given their close proximity and described how SHAREdash fills this gap since states do not typically share data with one another.

> *Being able to look at data kind of in the same place and say 'Oh, what does your county look like?' You know, which borders us in Oregon, but borders like three of the counties that I oversee. So, what's happening in their county? I can look that up and see if we're seeing similar trends, and the three counties that border that county. So having that originality, I think, is great and is probably a reason that I would go to the website to look at that at some point, or my team would.* [participant 9]

> *It is nice that it includes multiple states, because we are border county in our state, and so a lot of times things that we see are only for Oregon. But we're right next to a couple of Washington counties and it would be great to, you know, compare in that manner as well…It's always really helpful when we can look at, you know, what is our information compared to our neighboring counties, what does our data look like compared to counties of similar size.* [participant 2]

> *I like that you can see a big picture, regionally. So not necessarily just like other counties in Idaho: being able to prepare to other regional and other states and perhaps similar geographic demographic areas that are comparable, but in different states kind of just to see what trends are like there comparatively.* [participant 4]

The third theme of "needs more up-to-date data" described the biggest challenge that participants identified to their overall satisfaction with SHAREdash.

> *I think just what I commented on already is the age of the data that is present. So, it is very difficult to make a decision on data that's extremely outdated. And it's hard to make it relevant to your case. And I know that data can be hard to gather and hard to access, but for those of us who are looking at data to make decisions, that complicates that entire scenario. You want us to use data to make decisions, we need good data to make those decisions. Somebody has to put the data out.* [participant 6]

Several participants acknowledged that none or out-of-date data are typical within public health, particularly for rural areas. "We're used to that, so I think for us that's not a missed expectation to click on it and be like 'Oh, there's not any new data.'…That for us, that's normal" [participant 2]. However, this is a clear barrier to satisfaction and future use of SHAREdash.

There were 2 themes identified on validity called "reputable data sources increases validity" and "impact of missing data decreases validity." Participants spoke about the second theme of "reputable data sources increases validity" by describing their confidence in the data quality and accuracy. A participant described this by stating that, "SHAREdash is a really amazing place to quickly get domestic violence rates across other states. And you can find the source easily. And it is a reputable source too" [participant 1].

Whereas, another participant emphasized more than just the high quality of the data sources, but also the fact that SHAREdash's team provided a second, external check on it:

> *So I think this is a great dashboard and it's so nice because part of my job is to pull [data] from all of these different data sources which I know SHARE has done, and it's been validated and checked and it's a combination of information from various places which is good to have.* [participant 3]

For the "impact of missing data decreases validity" theme, task 1 had participants refined the population of interest to their specific county, which for some participants resulted in SHAREdash indicating that there were no data for their respective county available. Participants described how missing data in the dashboard impacted their ability to completely address the tasks in the think-aloud evaluation and how it would impact their work.

XSL•FO
RenderX

> *I think one of my biggest challenges, and it tends to be a challenge everywhere not just like solely for the dashboards, is that a lot of times when there were things I wanted to look at and there wasn't any data available because our population isn't that big.* [participant 8]

Despite multiple participants acknowledging that problems with data availability for rural areas is a known issue and is not a fault of the dashboard, they still expressed frustration and dissatisfaction about this issue.

> *I was bummed when it didn't have the data that I was looking for. But like I said, it's probably just a result of that data not being available.* [participant 4]

Discussion

Principal Findings

This SDLC stage 3 usability evaluation of SHAREdash, a dashboard designed for rural public health, indicates that overall SHAREdash is an efficient and valid tool that users reported being satisfied with. Task analyses and qualitative findings illustrate how SHAREdash's collaborative co-design process resulted in a tool that is easy to use and supports rural public health professionals' work. Thematic results also identified areas where SHAREdash can be improved to increase its usability such as changing some of the terms and names used and considering alternate ways for users to view and select information that are not just drop-down filters. However, this evaluation also uncovered usability issues related to the lack of public health data that go beyond design aspects and cannot be addressed through modifying SHAREdash's interface or navigation.

Issues related to obtaining quality public health data are well documented in the literature and include the critical problems of a lack of investment in public health data systems and infrastructure [36-39], issues with data quality [40-42] and data fragmentation [43,44], and the sparse data available about rural communities [8,45]. While every effort was made to include as much timely and comprehensive data as possible in SHAREdash, these larger data problems clearly impacted the usability of this tool. Thus, returning to the question posed by Thorpe and Gourevitch [5] regarding whether or not data dashboards for advancing health and equity are fulfilling their promise, findings from our study show that, to fully realize the potential of health equity–focused dashboards substantial investments in public health data need to be made. Unlike health care which benefited from the 2009 Health Information Technology for Economic and Clinical Health Act that supported the adoption of meaningful use of electronic health records [46,47], other systems, such as public health and social services, were not included in these incentives resulting in the numerous data issues seen today. The need for investments in public health data and systems was further clarified and reinforced by the COVID-19 pandemic which magnified many of these ongoing challenges [38,45,48]. To address this urgent problem, supportive policies that fund public health data collection and systems should leverage the successes of the Health Information Technology for Economic and Clinical Health Act and learn from the opportunities to address and alleviate these issues.

Another key finding from this study is also related to data. All the participants emphasized the significance of the trustworthiness of the data in SHAREdash. These results align with prior literature that has articulated the dual importance of dashboards to use data from reputable sources and clearly display or link to original data sources [49]. In a 2020 study by Young and Kitchin [50] that examined user perspectives of 4 different city's dashboards to create design guidelines, the authors stipulate how critical the veracity (eg, accuracy, source, and age) of the included data is. Our findings reinforce this work and indicate the utility of their design guidelines for creating data dashboards of municipal data. Future dashboards of municipal data should use the guidelines provided by Young and Kitchin [50] in the early design and development stages and work with target users to refine them for their specific project needs.

Satisfaction with SHAREdash was high with most participants describing the usefulness of the dashboard in supporting their work. Almost all the participants reported that they would like to make local-level comparisons that cross their respective states and articulated how difficult this currently is. Participants reported how comparisons between counties across different states can be more meaningful than within if they are able to filter for key factors such as population size or number of services available and how helpful it is that SHAREdash facilitates this easily. Enabling such comparisons points to the importance of aggregating large amounts of data across states, particularly for rural health departments that have unique needs and face different challenges than their urban counterparts [51]. It also indicates the importance of continuing to elucidate the unique needs of rural public health. Future research should focus on rural public health so that tailored design guidelines and specialized tools can be developed to support their work in addressing health disparities.

Our SDLC stage 3 usability assessment indicated that SHAREdash is meeting the goal of providing accurate, accessible, and relevant data via a user-centered dashboard to address health equity for rural communities. Next steps for SHAREdash will focus on identifying the elements key to its integration into LHDs using an implementation science approach that is outlined in stage 4 of the SDLC [30]. Planning for this phase is underway and is working closely with future end users to proactively identify and understand barriers to integration as this was a clear lesson learned from a similar study implementing an ICU dashboard [52]. Furthermore, investing in efforts to understand what is needed to support the uptake of health equity–focused dashboards in public health practice is critical to ensuring their impact [5] and aligns with previously identified public health research priorities [53-55] that highlights the importance of using implementation science to translate and assess innovations into public health practice to ensure reach. It is hoped that through the user-centered development and thoughtful translation of informatics, tools such as SHAREdash will address the existing health disparities and improve rural health equity.

XSL•FO
RenderX

Limitations

While our methods were rigorous, this study has limitations. Despite reaching data saturation, the sample size is small, consisting of all female-identifying participants, and limited to the northwest United States. Of note, the public health workforce is 79% women [56], which made diversity by sex difficult to obtain. Future studies would benefit from a larger and broader sample. Additionally, the think-aloud task analysis did not have a control arm where participants completed the tasks without SHAREdash to provide a comparison. Finally, while aligned with the SLC stage 3 usability evaluation components, this study did not examine other aspects of usability that are outlined in the literature, and thus, might have missed certain usability aspects [57].

Conclusions

Evaluating the usability of health equity dashboards is crucial to creating effective and valuable tools. Our findings indicate that SHAREdash, a public health dashboard created to support promoting health equity among rural communities, is an efficient, valid tool that overall users are satisfied with. Results strongly suggest that the utility of dashboards such as SHAREdash would be improved with the availability of more public health data and supportive policies to achieve robust collection of public health data would be beneficial. Future research should continue to focus on building tools that meet the unique needs of professionals working in rural public health to better support and equip them to alleviate rural health disparities.

Acknowledgments

This work was funded by the Department of Health and Human Services, Office of the Assistant Secretary for Health (grant #1 CPIMP171144-01-00). We would like to thank the public health professionals who participated in this study; the SHARE-NW team members who are not authors of this paper for their contributions to this project; and the SHARE-NW Equity Advisory Committee members for their expertise and guidance.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Usability evaluation types and goals based on the system development lifecycle (SDLC) stage (adapted from Yen and Bakken [30], which is published under Creative Commons Attribution-NonCommercial-NoDerivs licence).
[DOCX File , 19 KB - humanfactors_v11i1e51666_app1.docx]

Multimedia Appendix 2
Information on SHAREdash data sources.
[DOCX File , 34 KB - humanfactors_v11i1e51666_app2.docx]

Multimedia Appendix 3
Information on task analyses.
[DOCX File , 18 KB - humanfactors_v11i1e51666_app3.docx]

Multimedia Appendix 4
Semistructured interview guide.
[DOCX File , 20 KB - humanfactors_v11i1e51666_app4.docx]

References

1. Dasgupta N, Kapadia F. The future of the public health data dashboard. Am J Public Health 2022;112(6):886-888 [FREE Full text] [doi: 10.2105/AJPH.2022.306871] [Medline: 35613427]
2. Vahedi A, Moghaddasi H, Asadi F, Hosseini AS, Nazemi E. Applications, features and key indicators for the development of COVID-19 dashboards: a systematic review study. Inform Med Unlocked 2022;30:100910 [FREE Full text] [doi: 10.1016/j.imu.2022.100910] [Medline: 35342788]
3. Carroll LN, Au AP, Detwiler LT, Fu TC, Painter IS, Abernethy NF. Visualization and analytics tools for infectious disease epidemiology: a systematic review. J Biomed Inform 2014;51:287-298 [FREE Full text] [doi: 10.1016/j.jbi.2014.04.006] [Medline: 24747356]
4. Murphy DR, Savoy A, Satterly T, Sittig DF, Singh H. Dashboards for visual display of patient safety data: a systematic review. BMJ Health Care Inform 2021;28(1):e100437 [FREE Full text] [doi: 10.1136/bmjhci-2021-100437] [Medline: 34615664]
5. Thorpe LE, Gourevitch MN. Data dashboards for advancing health and equity: proving their promise? Am J Public Health 2022;112(6):889-892 [FREE Full text] [doi: 10.2105/AJPH.2022.306847] [Medline: 35446603]

XSL•FO
RenderX

6. The foundational public health services. Public Health Accreditation Board. 2022. URL: https://phaboard.org/center-for-innovation/public-health-frameworks/the-foundational-public-health-services [accessed 2024-03-01]
7. Gibson PJ, Shah GH, Streichert LC, Verchick L. Urgent challenges for local public health informatics. J Public Health Manag Pract 2016;22(Suppl 6):S6-S8 [FREE Full text] [doi: 10.1097/PHH.0000000000000479] [Medline: 27684620]
8. Harris JK, Beatty K, Leider JP, Knudson A, Anderson BL, Meit M. The double disparity facing rural local health departments. Annu Rev Public Health 2016;37:167-184 [FREE Full text] [doi: 10.1146/annurev-publhealth-031914-122755] [Medline: 26735428]
9. Bekemeier B, Marlowe J, Squires LS, Tebaldi J, Park S. Perceived need versus current spending: gaps in providing foundational public health services in communities. J Public Health Manag Pract 2018;24(3):271-280. [doi: 10.1097/PHH.0000000000000612] [Medline: 28832431]
10. Bekemeier B, Park S, Backonja U, Ornelas I, Turner AM. Data, capacity-building, and training needs to address rural health inequities in the Northwest United States: a qualitative study. J Am Med Inform Assoc 2019;26(8-9):825-834 [FREE Full text] [doi: 10.1093/jamia/ocz037] [Medline: 30990561]
11. Bekemeier B, Pantazis A, Yip MPY, Kwan-Gett T. Developing the evidence for public health systems to battle vaccine preventable disease at the local level: data challenges and strategies for advancing research. J Public Health Manag Pract 2017;23(2):131-137. [doi: 10.1097/PHH.0000000000000411] [Medline: 27798522]
12. Burgard SA, Chen PV. Challenges of health measurement in studies of health disparities. Soc Sci Med 2014;106:143-150 [FREE Full text] [doi: 10.1016/j.socscimed.2014.01.045] [Medline: 24561776]
13. Feng W, Martin EG. Fighting obesity at the local level? an analysis of predictors of local health departments' policy involvement. Prev Med 2020;133:106006. [doi: 10.1016/j.ypmed.2020.106006] [Medline: 32007526]
14. Zhang X, Luo H, Gregg EW, Mukhtar Q, Rivera M, Barker L, et al. Obesity prevention and diabetes screening at local health departments. Am J Public Health 2010;100(8):1434-1441 [FREE Full text] [doi: 10.2105/AJPH.2009.168831] [Medline: 20558810]
15. Mays GP, Smith SA, Ingram RC, Racster LJ, Lamberth CD, Lovely ES. Public health delivery systems: evidence, uncertainty, and emerging research needs. Am J Prev Med 2009;36(3):256-265. [doi: 10.1016/j.amepre.2008.11.008] [Medline: 19215851]
16. Mete C, Cioffi JP, Lichtveld MY. Are public health services available where they are most needed? an examination of local health department services. J Public Health Manag Pract 2003;9(3):214-223. [doi: 10.1097/00124784-200305000-00006] [Medline: 12747318]
17. Santerre RE. Jurisdiction size and local public health spending. Health Serv Res 2009;44(6):2148-2166 [FREE Full text] [doi: 10.1111/j.1475-6773.2009.01006.x] [Medline: 19656226]
18. Munzner T. A nested model for visualization design and validation. IEEE Trans Vis Comput Graph 2009;15(6):921-928. [doi: 10.1109/TVCG.2009.111] [Medline: 19834155]
19. Backonja U, Park S, Kurre A, Yudelman H, Heindel S, Schultz M, et al. Supporting rural public health practice to address local-level social determinants of health across Northwest states: development of an interactive visualization dashboard. J Biomed Inform 2022;129:104051 [FREE Full text] [doi: 10.1016/j.jbi.2022.104051] [Medline: 35318149]
20. Abran A, Khelifi A, Suryn W, Seffah A. Usability meanings and interpretations in ISO standards. Softw Qual J 2003;11:325-338 [FREE Full text] [doi: 10.1023/a:1025869312943]
21. Bevan N, Carter J, Earthy J, Geis T, Harker S. New ISO standards for usability, usability reports and usability measures. In: Kurosu M, editor. Human-Computer Interaction: Theory, Design, Development and Practice. HCI 2016. Lecture Notes in Computer Science Vol. 9731. Cham: Springer; 2016:268-278.
22. Kushniruk AW, Triola MM, Borycki EM, Stein B, Kannry JL. Technology induced error and usability: the relationship between usability problems and prescription errors when using a handheld application. Int J Med Inform 2005;74(7-8):519-526. [doi: 10.1016/j.ijmedinf.2005.01.003] [Medline: 16043081]
23. Carayon P, Du S, Brown R, Cartmill R, Johnson M, Wetterneck TB. EHR-related medication errors in two ICUs. J Healthc Risk Manag 2017;36(3):6-15 [FREE Full text] [doi: 10.1002/jhrm.21259] [Medline: 28099789]
24. Brown CL, Mulcaster HL, Triffitt KL, Sittig DF, Ash JS, Reygate K, et al. A systematic review of the types and causes of prescribing errors generated from using computerized provider order entry systems in primary and secondary care. J Am Med Inform Assoc 2017;24(2):432-440 [FREE Full text] [doi: 10.1093/jamia/ocw119] [Medline: 27582471]
25. Gephart S, Carrington JM, Finley B. A systematic review of nurses' experiences with unintended consequences when using the electronic health record. Nurs Adm Q 2015;39(4):345-356. [doi: 10.1097/NAQ.0000000000000119] [Medline: 26340247]
26. Schnall R, Mosley JP, Iribarren SJ, Bakken S, Carballo-Diéguez A, Iii WB. Comparison of a user-centered design, self-management app to existing mHealth apps for persons living with HIV. JMIR Mhealth Uhealth 2015;3(3):e91 [FREE Full text] [doi: 10.2196/mhealth.4882] [Medline: 26385783]
27. McCurdie T, Taneva S, Casselman M, Yeung M, McDaniel C, Ho W, et al. mHealth consumer apps: the case for user-centered design. Biomed Instrum Technol 2012;Suppl:49-56 [FREE Full text] [doi: 10.2345/0899-8205-46.s2.49] [Medline: 23039777]
28. Eysenbach G. The law of attrition. J Med Internet Res 2005;7(1):e11 [FREE Full text] [doi: 10.2196/jmir.7.1.e11] [Medline: 15829473]

XSL•FO
RenderX

29. Kaufman D, Roberts WD, Merrill J, Lai TY, Bakken S. Applying an evaluation framework for health information system design, development, and implementation. Nurs Res 2006;55(Suppl 2):S37-S42. [doi: 10.1097/00006199-200603001-00007] [Medline: 16601633]
30. Yen PY, Bakken S. Review of health information technology usability study methodologies. J Am Med Inform Assoc 2012;19(3):413-422 [FREE Full text] [doi: 10.1136/amiajnl-2010-000020] [Medline: 21828224]
31. Lewis C. Using the "thinking-aloud" method in cognitive interface design. IBM T. J. Watson Research Center. 1982. URL: https://dominoweb.draco.res.ibm.com/2513e349e05372cc852574ec0051eea4.html#:~:text=%22Thinking%2Daloud%22%20is%20a,to%20use%20a%20computer%20system [accessed 2024-05-03]
32. Maramba I, Chatterjee A, Newman C. Methods of usability testing in the development of eHealth applications: a scoping review. Int J Med Inform 2019;126:95-104. [doi: 10.1016/j.ijmedinf.2019.03.018] [Medline: 31029270]
33. Nielsen J. Usability Engineering. San Francisco, CA: Morgan Kaufmann; 1994.
34. Bevan N, Carter J, Harker S. ISO 9241-11 revised: what have we learnt about usability since 1998? In: Kuroso M, editor. Human-Computer Interaction: Design and Evaluation. HCI 2015. Lecture Notes in Computer Science Vol. 9169. Cham: Springer; 2015:143-151.
35. Braun V, Clarke V. Using thematic analysis in psychology. Qual Res Psychol 2006;3(2):77-101. [doi: 10.1191/1478088706qp063oa]
36. Alfonso YN, Leider JP, Resnick B, McCullough JM, Bishai D. US public health neglected: flat or declining spending left states ill equipped to respond to COVID-19. Health Aff (Millwood) 2021;40(4):664-671 [FREE Full text] [doi: 10.1377/hlthaff.2020.01084] [Medline: 33764801]
37. Orr JM, Leider JP, Singh S, Balio CP, Yeager VA, Bekemeier B, et al. Regarding investment in a healthier future: impact of the 2012 Institute of Medicine finance report. J Public Health Manag Pract 2022;28(1):E316-E323. [doi: 10.1097/PHH.0000000000001209] [Medline: 32956294]
38. Degutis LC, Shoaf K, Aragón TJ, Atchison C, Dyjack D, Hites L, et al. The urgent need for public health preparedness funding and support. Am J Public Health 2021;111(3):363-365. [doi: 10.2105/AJPH.2020.306068] [Medline: 33566658]
39. Yeager VA, Balio CP, McCullough JM, Leider JP, Orr J, Singh SR, et al. Funding public health: achievements and challenges in public health financing since the Institute of Medicine's 2012 report. J Public Health Manag Pract 2022;28(1):E244-E255. [doi: 10.1097/PHH.0000000000001283] [Medline: 33605671]
40. Dolley S. Big data's role in precision public health. Front Public Health 2018;6:68 [FREE Full text] [doi: 10.3389/fpubh.2018.00068] [Medline: 29594091]
41. Khoury MJ, Engelgau M, Chambers DA, Mensah GA. Beyond public health genomics: can big data and predictive analytics deliver precision public health? Public Health Genomics 2018;21(5-6):244-250 [FREE Full text] [doi: 10.1159/000501465] [Medline: 31315115]
42. Hoffman S, Podgurski A. Big bad data: law, public health, and biomedical databases. J Law Med Ethics 2013;41(Suppl 1):56-60. [doi: 10.1111/jlme.12040] [Medline: 23590742]
43. Velmovitsky PE, Bevilacqua T, Alencar P, Cowan D, Morita PP. Convergence of precision medicine and public health into precision public health: toward a big data perspective. Front Public Health 2021;9:561873 [FREE Full text] [doi: 10.3389/fpubh.2021.561873] [Medline: 33889555]
44. Iyamu I, Gómez-Ramírez O, Xu AX, Chang HJ, Watt S, Mckee G, et al. Challenges in the development of digital public health interventions and mapped solutions: findings from a scoping review. Digit Health 2022;8:20552076221102255 [FREE Full text] [doi: 10.1177/20552076221102255] [Medline: 35656283]
45. Bekemeier B, Heitkemper E, Backonja U, Whitman G, Schultz M, Jiang Y, et al. Rural public health data challenges during the COVID-19 pandemic: the case for building better systems ahead of a public health crisis. J Public Health Manag Pract 2023;29(4):496-502. [doi: 10.1097/PHH.0000000000001726] [Medline: 36867496]
46. Health IT legislation. HealthIT.gov. 2023. URL: https://www.healthit.gov/topic/laws-regulation-and-policy/health-it-legislation [accessed 2023-08-02]
47. Burde H. Health law the hitech act-an overview. Virtual Mentor 2011;13(3):172-175 [FREE Full text] [doi: 10.1001/virtualmentor.2011.13.3.hlaw1-1103] [Medline: 23127320]
48. Budd J, Miller BS, Manning EM, Lampos V, Zhuang M, Edelstein M, et al. Digital technologies in the public-health response to COVID-19. Nat Med 2020;26(8):1183-1192 [FREE Full text] [doi: 10.1038/s41591-020-1011-4] [Medline: 32770165]
49. Ansari B, Martin EG. Development of a usability checklist for public health dashboards to identify violations of usability principles. J Am Med Inform Assoc 2022;29(11):1847-1858. [doi: 10.1093/jamia/ocac140] [Medline: 35976140]
50. Young GW, Kitchin R. Creating design guidelines for building city dashboards from a user's perspectives. Int J Hum Comput Stud 2020;140:102429 [FREE Full text] [doi: 10.1016/j.ijhcs.2020.102429]
51. Leider JP, Meit M, McCullough JM, Resnick B, Dekker D, Alfonso YN, et al. The state of rural public health: enduring needs in a new decade. Am J Public Health 2020;110(9):1283-1290 [FREE Full text] [doi: 10.2105/AJPH.2020.305728] [Medline: 32673103]

XSL•FO
RenderX

52. Wac M, Craddock I, Chantziara S, Campbell T, Santos-Rodriguez R, Davidson B, et al. Design and evaluation of an intensive care unit dashboard built in response to the COVID-19 pandemic: semistructured interview study. JMIR Hum Factors 2023;10:e49438 [FREE Full text] [doi: 10.2196/49438] [Medline: 37751239]
53. Shelton RC, Cooper BR, Stirman SW. The sustainability of evidence-based interventions and practices in public health and health care. Annu Rev Public Health 2018;39:55-76 [FREE Full text] [doi: 10.1146/annurev-publhealth-040617-014731] [Medline: 29328872]
54. Chambers DA. Commentary: increasing the connectivity between implementation science and public health: advancing methodology, evidence integration, and sustainability. Annu Rev Public Health 2018;39:1-4. [doi: 10.1146/annurev-publhealth-110717-045850] [Medline: 29272164]
55. Estabrooks PA, Brownson RC, Pronk NP. Dissemination and implementation science for public health professionals: an overview and call to action. Prev Chronic Dis 2018;15:E162 [FREE Full text] [doi: 10.5888/pcd15.180525] [Medline: 30576272]
56. Nicolaus T. Public health addressing gender gap in workforce leadership. Nation's Health 2022;52(6):1-6 [FREE Full text]
57. Nielsen J. Usability 101: introduction to usability. NN/g Nielsen Norman Group. 2013. URL: https://www.nngroup.com/articles/usability-101-introduction-to-usability [accessed 2023-08-02]

Abbreviations

IRB: institutional review board
LHD: local health department
SDLC: system development lifecycle
SHAREdash: solutions in health analytics for rural equity across the northwest dashboard
SHARE-NW: solutions in health analytics for rural equity across the northwest

Edited by D Gooch; submitted 07.08.23; peer-reviewed by CH Lai, Anonymous; comments to author 02.01.24; revised version received 24.03.24; accepted 18.04.24; published 05.06.24.

Please cite as:
Heitkemper E, Hulse S, Bekemeier B, Schultz M, Whitman G, Turner AM
The Solutions in Health Analytics for Rural Equity Across the Northwest (SHARE-NW) Dashboard for Health Equity in Rural Public Health: Usability Evaluation
JMIR Hum Factors 2024;11:e51666
URL: https://humanfactors.jmir.org/2024/1/e51666
doi:10.2196/51666
PMID:38837192

©Elizabeth Heitkemper, Scott Hulse, Betty Bekemeier, Melinda Schultz, Greg Whitman, Anne M Turner. Originally published in JMIR Human Factors (https://humanfactors.jmir.org), 05.06.2024. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on https://humanfactors.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Implementing a Hospital Call Center Service for Mental Health in Uganda: User-Centered Design Approach

Johnblack K Kabukye[1,2], MBChB, MSc, PhD; Rosemary Namagembe[3], BSN; Juliet Nakku[4], MBChB, MMed; Vincent Kiberu[5], PhD; Marie Sjölinder[6], PhD; Susanne Nilsson[7], PhD; Caroline Wamala-Larsson[1], PhD

[1]SPIDER - The Swedish Program for ICT in Developing Regions, Department of Computer and Systems Sciences, Stockholm University, Stockholm, Sweden

[2]Uganda Cancer Institute, Kampala, Uganda

[3]Hutchinson Centre Research Institute of Uganda, Uganda Cancer Institute, Kampala, Uganda

[4]Butabika National Referral and Teaching Mental Hospital, Kampala, Uganda

[5]School of Public Health, Makerere University, Kampala, Uganda

[6]RISE - Research Institutes of Sweden, Kista, Sweden

[7]Unit for Integrated Product Development and Design, Department of Machine Design, KTH Royal Institute of Technology, Stockholm, Sweden

Corresponding Author:
Johnblack K Kabukye, MBChB, MSc, PhD
SPIDER - The Swedish Program for ICT in Developing Regions
Department of Computer and Systems Sciences
Stockholm University
Borgarfjordsgatan 12 (NOD Building)
PO Box 7003 Kista
Stockholm, SE-164 07
Sweden
Phone: 46 8162000
Fax: 46 87039025
Email: kabukye@dsv.su.se

Abstract

Background: Mental health conditions are a significant public health problem globally, responsible for >8 million deaths per year. In addition, they lead to lost productivity, exacerbate physical illness, and are associated with stigma and human rights violations. Uganda, like many low- and middle-income countries, faces a massive treatment gap for mental health conditions, and numerous sociocultural challenges exacerbate the burden of mental health conditions.

Objective: This study aims to describe the development and formative evaluation of a digital health intervention for improving access to mental health care in Uganda.

Methods: This qualitative study used user-centered design and design science research principles. Stakeholders, including patients, caregivers, mental health care providers, and implementation experts (N=65), participated in focus group discussions in which we explored participants' experience of mental illness and mental health care, experience with digital interventions, and opinions about a proposed digital mental health service. Data were analyzed using the Consolidated Framework for Implementation Research to derive requirements for the digital solution, which was iteratively cocreated with users and piloted.

Results: Several challenges were identified, including a severe shortage of mental health facilities, unmet mental health information needs, heavy burden of caregiving, financial challenges, stigma, and negative beliefs related to mental health. Participants' enthusiasm about digital solutions as a feasible, acceptable, and convenient method for accessing mental health services was also revealed, along with recommendations to make the service user-friendly, affordable, and available 24×7 and to ensure anonymity. A hospital call center service was developed to provide mental health information and advice in 2 languages through interactive voice response and live calls with health care professionals and peer support workers (recovering patients). In the 4 months after launch, 456 calls, from 236 unique numbers, were made to the system, of which 99 (21.7%) calls went to voicemails (out-of-office hours). Of the remaining 357 calls, 80 (22.4%) calls stopped at the interactive voice response, 231 (64.7%) calls were answered by call agents, and 22 (6.2%) calls were not answered. User feedback was positive, with callers appreciating the inclusion of peer support workers who share their recovery journeys. However, some participant recommendations

XSL•FO
RenderX

(eg, adding video call options) or individualized needs (eg, prescriptions) could not be accommodated due to resource limitations or technical feasibility.

Conclusions: This study demonstrates a systematic and theory-driven approach to developing contextually appropriate digital solutions for improving mental health care in Uganda and similar contexts. The positive reception of the implemented service underscores its potential impact. Future research should address the identified limitations and evaluate clinical outcomes of long-term adoption.

(JMIR Hum Factors 2024;11:e53976) doi:10.2196/53976

KEYWORDS

mHealth; mobile health; digital health; digital solution; digital solutions; digital intervention; digital interventions; mental health; awareness; Uganda; Africa; African; user centred; user centered; design; qualitative; focus group; focus groups; call centre; call centres; call center; call centers; mental; experience; experiences; attitude; attitudes; opinion; perception; perceptions; perspective; perspectives; cocreated; cocreation; service; services; mobile phone

Introduction

Mental health conditions are an important public health issue globally, responsible for >8 million deaths per year [1-5]. Three million people die annually from the harmful use of alcohol, and 1 person dies every 40 seconds by suicide [1,2]. An estimated 970.1 million people (12.6% of the global population) experience some form of mental health problem [4]. Mental health conditions account for 5% of the global disability-adjusted life years and 12% to 20% of years lived with disability [4,6]. People with mental health conditions, on average, die 20 years prematurely [4,6] both due to mental as well as physical illnesses because mental health conditions are a risk factor for, or can complicate, physical illnesses, including physical injury and road traffic accidents, HIV or AIDS, cardiovascular diseases, and cancer [5,7]. People with mental health conditions also experience severe human rights violations, stigma, discrimination, abuse, and generally poor socioeconomic status [5,7-9].

Unfortunately, >75% of people with mental health problems do not have access to the care they need [1-3]. This is especially true for Uganda [10,11] and similar low- and middle-income countries (LMICs) where the treatment gap for mental disorders reaches 90% [12-14]. It is estimated that the ratio of mental health workers to population is 200 times smaller in LMICs compared with the high-income countries [3]. In LMICs, mental health is underprioritized in the face of other competing public health challenges such as HIV and AIDS, tuberculosis, malaria, and maternal and child health. Uganda, for example, spends 9.8% of its gross domestic product on health care, but <1% of this goes toward mental health care [10,11]. Consequently, Uganda experiences a shortage of mental health care facilities and professionals and poor and inconsistent access to medication and related mental health services [11]. In addition, most of the health workforce is limited to urban areas, yet >80% of the population lives in rural areas, thus geographically isolated from even the limited care available. Other important challenges facing mental health in Uganda include social norms [15], beliefs (such as witchcraft), lack of awareness of mental health disorders [8,11,16], pervasive stigma, and sociopolitical conflicts [13,17]; these not only result in an increase in the incidence of mental health problems but also lead to many people with mental health problems not seeking care and going undiagnosed.

To address some of the abovementioned challenges and improve access to mental health services in Uganda, we implemented the project *digitalizing mental health care and access in Uganda*. In this project, we followed a user-centered design (UCD) and a cocreation process to set up a hospital call center service to provide mental health information and advice to patients, caregivers, and the general public. This paper aimed to describe the development and formative evaluation of this mental health call center service.

Methods

Ethical Considerations

Ethical approval for the research study was obtained from the Makerere University School of Public Health research ethics committee (#SPH-2021-153) and the Uganda National Council of Science and Technology (#HS1868ES). All participants provided written informed consent before participating in the study activities.

Study Design

We conducted a qualitative case study using the principles of UCD [18-20] and design science research (DSR) [21,22]. UCD focuses on understanding and prioritizing the needs, preferences, and behavior of end users of a product throughout its development life cycle. UCD, therefore, calls for iterative and collaborative engagement of users to ensure high usability and utility of the product. DSR is a structured approach to creating and evaluating innovative solutions or artifacts, where the design process is treated as research that contributes to knowledge for improving the functional performance of artifacts. The steps involved in DSR mirror UCD and include the following: (1) identifying the problem and motivation (understanding user experiences and context of use); (2) defining the objective of the solution (specifying the requirements); (3) designing and development of (often novel) solutions using participatory or cocreation processes; and (4) demonstrating and evaluating the solutions to validate against requirements, assess usability, and long-term adoption. These steps help identify the facilitators and barriers of adoption so that they can be addressed early on in the project life cycle, allow user engagement and facilitate buy-in, ensure that the product fits the context of use and purpose, and has good usability [23,24] and clinical utility [25].

XSL•FO
RenderX

In the following sections, we describe each of the above 4 steps. Note that there was overlap and iterations over the steps as per the UCD best practice. To ease readability, we report the procedure and results from each step. Thereafter, we provide a general discussion and conclusion.

Step 1: Understanding User Experiences and Context to Identify the Problems

Participants and Recruitment

The participants included adults (≥18 years), patients recovering from mental disorders, caregivers of such patients, peer support workers (PSWs), mental health care providers, and persons involved in the implementation of call centers for telecoms or other health care centers. The health care providers, patients, caregivers, and PSWs were recruited from the Butabika National Mental Referral Hospital in Kampala, Uganda, which is also the site of implementation. Sampling was purposive to include different cadres and expertise of providers (informed by the third author, who is the head of Butabika Hospital) and to represent different mental health conditions, levels of education, and socioeconomic status of patients and caregivers to get diversity of experiences and views. The investigators (JKK, JN, and VK) physically approached the health care providers at Butabika Hospital, explained the project's purpose and the research activities involved (including participation in multiple group discussions and workshops), and obtained consent from those interested. These health care providers then reached out to patients under their care, caregivers, and PSWs; provided them with information about the study; and invited those interested for consent by the investigators, who explained the participants' rights and voluntary nature of participation. Participants in the last stakeholder category were recruited through the network of the first author who works in the digital health field in Uganda.

Data Collection

We conducted semistructured focus group discussions (FGDs) in which we explored participants' experience of mental illness and mental health care in Uganda (including unmet information and supportive care needs); experience with call center services from the commercial service sector or other digital health care services; and opinions about a proposed digital mental health service (ie, feasibility, appropriateness, expected benefits, or recommendations for successful implementation). The FGD guide is shown in Textbox 1. There was flexibility in the order of the probes to allow free flow of ideas, with additional probes for clarification added by moderators as issues of interest arose. In addition, certain issues or probes were discussed in detail, paraphrased, or left out as appropriate depending on relevance to the session participant or if such a topic had been sufficiently explored in the prior sessions.

The sessions were conducted in English and Luganda (the lingua franca in Uganda) as appropriate for the participant category. In addition, we held male-only and female-only FGD sessions for patients to reduce the possibility that some participants would overshadow others during the discussion, but other sessions were mixed to ensure rich discussions since diverse viewpoints from different participants inspire others and spur discussion. The first and fourth authors (JKK and VK) were the moderators, while the second author (RN) was a notetaker. The sessions were audio recorded and later transcribed by the second author, who also translated the sessions in Luganda into English for analysis. The FGDs took place in November 2021.

We drew on the Consolidated Framework for Implementation Research (CFIR) [26] to inform data collection (and analysis; see the *Data Analysis* section). The CFIR is a *metatheoretical framework* developed by consolidating several implementation science theories into one comprehensive taxonomy of clearly defined, nonoverlapping constructs related to disseminating and implementing evidence-based interventions. These constructs fall into five domains: (1) the individuals affected or involved in the implementation, (2) the innovation (intervention), (3) the inner setting (organization) where the innovation is implemented, (4) the outer setting (wider societal context), and (5) the process of implementation. The CFIR is one of the most widely used theoretical frameworks to identify implementation barriers and facilitators (ie, a determinant framework) [27-31]. Following a literature review and feedback from researchers who have used the CFIR [29], a recent update, dubbed "CFIR 2.0," has been made, in addition to a CFIR outcomes addendum [32]. These updates have provided further clarification between constructs, including, for example, a distinction between "implementation determinants," which relate to the context, versus "innovation determinants," which relate to the characteristics of the innovation (eg, ease of use, relative advantage, cost, and efficacy of a technology). The implementation and innovation determinants inform the implementation process (needs assessment, user engagement, tailoring to user needs, incentives, marketing, etc) and moderate the anticipated and actual implementation outcomes through the antecedent assessments (tension or readiness for change, feasibility, acceptability, appropriateness, etc). The updates to the CFIR make it also useful for informing the design, implementation, and evaluation of innovations (ie, a process and evaluation framework) [30,31]. Figure 1 shows an adaptation of the CFIR and its recent updates as used in this study [26,29,32].

XSL•FO
RenderX

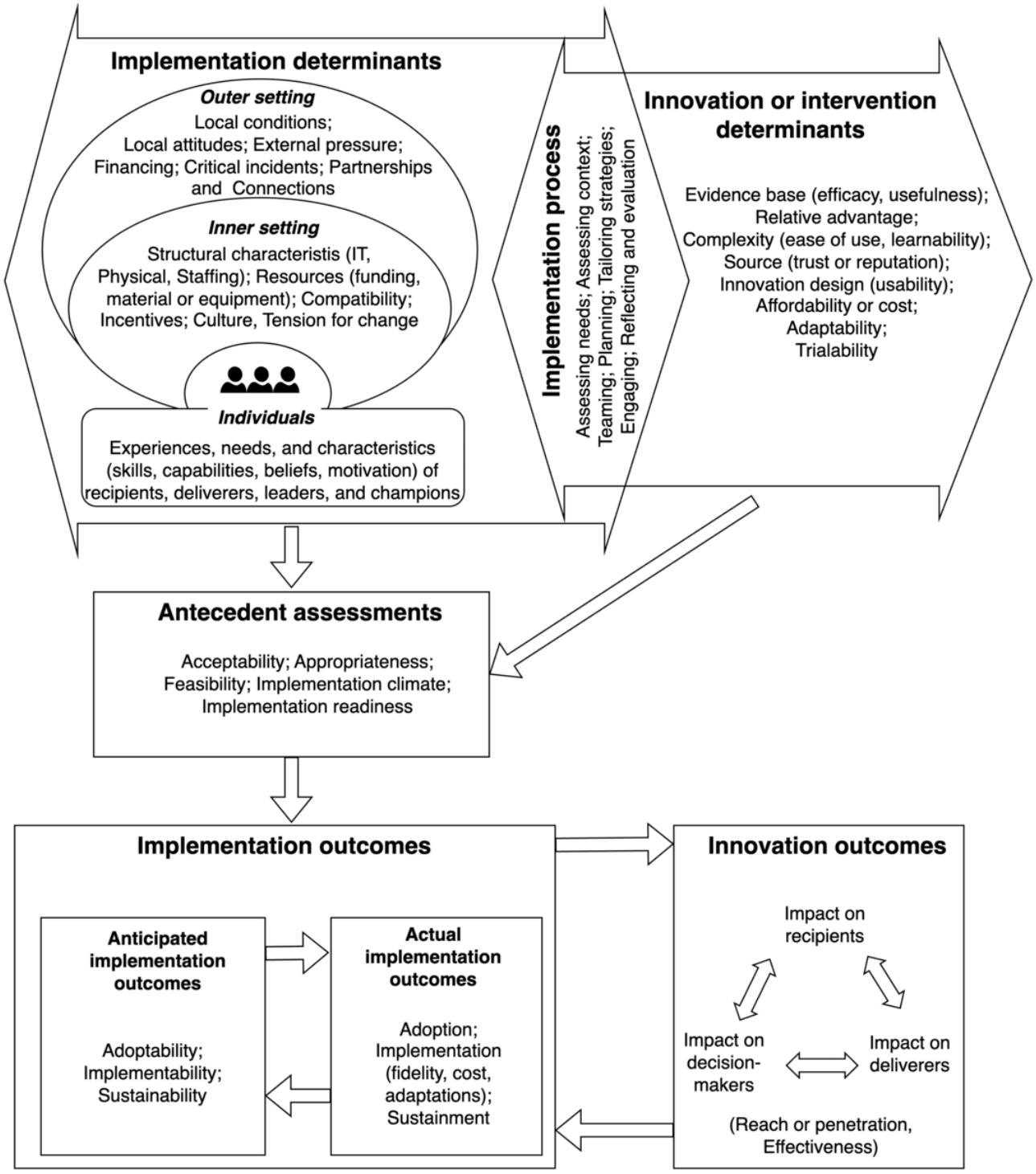
Textbox 1. Focus group discussion guide.

Topic and questions or probes

- Participants' understanding of mental health and mental problems
  - What is mental health, and what is mental health illness?
  - Do you know any forms of mental illness? What are the signs and symptoms?
  - What do you do with a person who has mental problems? What have you experienced? What is usually done, and what should be the correct thing to do?
  - Where can one get treatment? Probe about alternative healers, witchcraft, religious healers, etc
  - How are mental health or mental illnesses viewed in your community? Probe about stigma, myths, fear, and marginalization
- Mental health information, psychoeducation, and psychosocial support
  - What information about mental health problems or mental health care do you wish you knew early on in your mental illness journey?
  - What issues or topics do you think are the most important to address now? Are there any topics or issues that you still need information about? Give examples.
  - How or where do you get information about mental health and mental illnesses? Which ones are the best or preferred?
  - Tell us any challenges or limitations of these information sources.
  - Are there any services or persons that support you to cope with mental illness or care for your relatives with mental illness? Tell us more about these.
- Telemental health services
  - Tell us about your experience with interactive voice response (IVR) system or call centers: which industry or business? Any challenges and advantages? (moderator to explain IVR if participants do not know and can use the examples of telecoms or bank customer care lines to explain)
  - What are your thoughts on using such IVR systems for mental health information and care (telemental service)? Probes any experience of telemental health services, anticipated benefits, limitations, considerations on how to make it work, concerns about timing, phone ownership and access, privacy, etc. Probe for details and examples.
  - What are the likely barriers or facilitators for such a service?
  - Any thoughts about staffing and the role of peer support workers (PSWs)? Probe about acceptability to patients, benefit to PSWs, any anticipated challenges, and how to mitigate them.
  - Any other thoughts about using technology in mental health care?

XSL•FO
RenderX

Figure 1. The Consolidated Framework for Implementation Research (CFIR) used in this study [26,29,32].



Data Analysis

A directed (deductive) content analysis approach [33] was used. We began with a rapid qualitative analysis [34-36] of the FGDs in order to quickly identify the requirements and other insights needed to inform initial iterations of system development (see *Step 2: Specifying the Requirements of the System* section). Rapid qualitative analysis is aimed at getting actionable and targeted insights in a timely manner and is suitable for studies such as this one, where there is a need to refine and adapt an intervention or program, as opposed to developing new theories. A deductive approach is taken, using existing theories or frameworks (in our case, the CFIR) to summarize the qualitative data into, for example, intervention characteristics or barriers and facilitators. In rapid qualitative research, data collection and analysis occur concurrently and iteratively, with findings from one phase informing the next iteration. The analysis is done not on the transcripts but on the summaries or notes taken during the FGDs or the audio recordings. In addition, multiple data collection methods are used to triangulate findings (eg, FGDs, field observations, debriefing, and reflections by the research team or other stakeholders and literature review). In our study, the rapid analysis was done by the first 4 authors (JKK, RN, JN, and VK) and involved note-taking during FGDs

XSL•FO
RenderX

and discussion and summarization of insights after each session. When necessary, recordings were listened to by researchers who, for example, was not present in the session before they contributed to the analysis or for validating the summaries. We summarized the findings into an initial list of mental health information topics to be covered by our system and design considerations (system requirements) based on the experiences and expectations of users, as well as contextual constraints.

Later, a traditional qualitative analysis was done [34-36]. The first and second authors (JKK and RN) independently read 1 of the 7 transcripts and extracted meaningful units or statements and coded them into themes related to the research objectives and the CFIR. They then met to discuss and refine the coding before independently coding the remaining transcripts. Three more meetings were held to compare and refine the coding, after which the findings were shared with all the authors for discussion and interpretation. We focused on saliency [37] rather than frequency of issues and codes, such that even if an issue was mentioned once or by 1 participant category, we coded it as long as it related to the research question and CFIR constructs. As such, we did not count or rank the codes and themes. Basic Office software (Microsoft Corp) was used for coding and summarizing the qualitative data.

Results

Participants and FGD Sessions

We conducted 7 FGDs, each with 8-10 participants, for a total of 65 participants. The participants were fairly balanced by sex (female participants: 35/65, 54%; male participants: 30/65, 46%), and their ages ranged from 21 to 64 years, with a median of 40, IQR 12 years). Each session lasted approximately 1.5 hours. Table 1 shows the details of the FGD sessions.

Table 1. Details of focus group discussion (FGD) sessions (N=65).

FGD session	Stakeholder category	Participants, n (%)	Sex	Language	Notes
1	Patients	10 (15)	Female	Luganda	Sessions in Luganda (the most commonly spoken local language) and separated males and females to ensure participants speak freely and not overshadowed by opposite sex. Diagnoses represented included bipolar affective disorder, schizophrenia, and psychosis.
2	Patients	8 (12)	Male	Luganda	__[a]
3	Caregivers	9 (14)	Mixed	English	Separate sessions in English and in Luganda to get opinions from participants of different education status (English is learned in school in Uganda and is proxy for education and socioeconomic status). Diagnoses represented included bipolar affective disorder, schizophrenia, psychosis, epilepsy, and alcohol and substance use disorder.
4	Caregivers	10 (15)	Mixed	Luganda	—
5	Health care providers	10 (15)	Mixed	English	Staff of Butabika Hospital involved in care for patients and community outreaches, including psychiatrists, psychologists, psychiatric nurses, and psychiatric clinical officers.
6	PSWs[b]	10 (15)	Mixed	English	Volunteers with lived experience of mental illness. They work with Butabika Hospital to share their personal experience and support and educate other patients. They receive small stipends from the hospital, patients they help, or projects and grants to facilitate their work. Diagnoses represented included bipolar affective disorder, schizophrenia, and psychosis.
7	Implementers	8 (12)	Mixed	English	Customer care for telecoms, developers of IVR[c] systems, and implementers of hospital call centers in HIV or AIDS and cancer, private telemedicine company (general care), and mental health NGOs[d].

[a]Not applicable.

[b]PSW: peer support worker.

[c]IVR: interactive voice response.

[d]NGO: nongovernmental organization.

Findings From the FGDs

Multimedia Appendix 1 shows the qualitative findings, including the CFIR domains, constructs, themes, and their explanation. Multimedia Appendix 2 contains illustrative quotes.

Overall, 39 themes emerged across 20 CFIR constructs in all the 5 domains and the antecedent assessments. The themes recurred across the participant groups, supporting their validity. The themes under the *individuals domain* highlighted several challenges that people with mental health conditions in Uganda

XSL•FO
RenderX

face, including the limited number of mental health care facilities, long distances to care, lack of mental health information, stigma against patients with mental health problems and their families, financial challenges, and unmet psychosocial needs. The themes also covered contextual issues that explain these challenges. These included issues about the nature of mental illness (chronic and with a high burden of caregiving); organizational issues (*inner setting*), such as understaffing of mental health facilities and frequent medication stockouts; and societal issues (*outer setting*), such as beliefs and cultural norms (eg, belief in witchcraft), which influence how people understand mental health problems and how they seek care. Themes under the domain *innovation determinants* covered participants' perception or expected benefit from the proposed mental health call center service, including affordability; familiarity with similar services and the technology (ubiquitous access to mobile phones); convenience; time and cost saving; and anonymity offered by telephone services, which protect users from the stigma. Finally, themes in the *implementation process* domain encompassed mostly participants' recommendations or strategies for successful implementation, such as linkage with other stakeholders involved in mental health care, marketing of the service (sensitization), training and supervision of staff for quality control, and the need to maintain the human touch rather than attempting to digitalize or automate mental health care delivery. These findings suggested the feasibility, acceptability, and appropriateness of the proposed solution (*antecedent assessments*).

There were also several insights or implicit findings not mentioned by the FGD participants but inferred from observations and the research team's understanding of the context. These are relevant for the implementation and can be mapped to CFIR constructs. For example, there has been an increase in the adoption of telemedicine in Uganda, especially following the COVID-19 pandemic, which has given credibility to such innovations and can explain the general enthusiasm shown by the participants (CFIR construct "evidence base" in *innovation determinants*). In fact, the participants in the implementers' category were themselves involved in implementing call centers for HIV or AIDS, private telemedicine clinics, and mental health NGOs and were aware of the growing scientific evidence globally that supports digital health. The COVID-19 pandemic is also an example of "critical incidents" that can disrupt (or encourage) implementation and delivery of innovations (*outer setting*) according to CFIR 2.0. Other issues included the external project grant (construct: "Financing"), Uganda government's positive digital transformation strategies and policies (construct: "External pressure"), and the position of Butabika Hospital as a national referral that is supposed to be exemplary (construct: "Performance measurement pressure").

Step 2: Specifying the Requirements of the System

Procedure and Team

Requirements were specified based on the understanding of the users' needs, challenges, and contextual constraints from the FGDs. The development team consisted of the first 3 authors (a physician and digital health expert, a research nurse, and a senior consultant psychiatrist, respectively), as well as a psychologist, a psychiatric nurse, and an IT professional specializing in telephone systems. The first 2 authors and the IT professional have previously worked together to set up a similar system at the Uganda Cancer Institute [38] from which they also drew insights. The team held 8 web-based meetings from December 2021 to March 2022 to iteratively discuss the system features, content (mental health information), and setup considerations. We started with the initial list from the rapid qualitative analysis (see the *Data Analysis* section), which we refined to remove conflicting requirements or those that are not feasible due to available resources (eg, video telemedicine). We also agreed on the priority features and mental health information topics.

Results

Multimedia Appendix 3 lists the high-level requirements and how they were addressed in the system design and implementation. The key of these requirements is that the system or the intervention provides correct mental health information and psychosocial support in a culturally sensitive and nonstigmatizing manner and in multiple languages. In addition, the system should be easy to use (navigate), accessible 24×7, and affordable (free) to users; there should be no long queues; and it should fit within the workflow of the staff and not increase their workload. Finally, it should ensure privacy and confidentiality to users' information, and risks of harm to users should be minimized through quality control measures, training, and professionalism of staff.

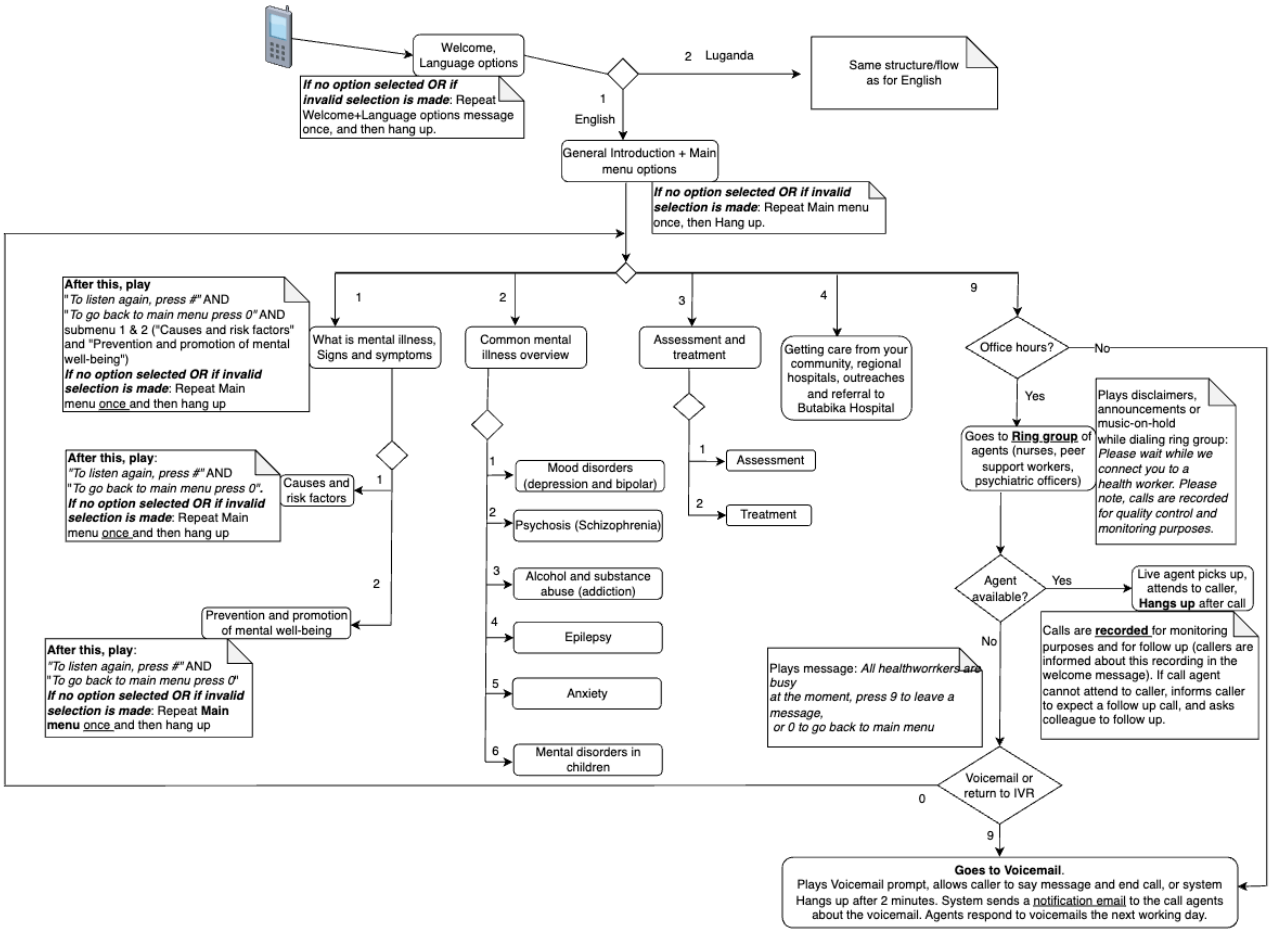
Step 3: Design and Development of the System

Procedure and Team

We designed and developed a telephone system for providing mental health information and advice to callers as per the requirements (Multimedia Appendix 3). The system consists of 3 complimentary components or features: an interactive voice response (IVR), live calls, and voicemails. The IVR is the first component that users interact with, and since it is automated, it is available 24×7. It contains mental health information in audio format in English and Luganda. Callers get navigation instructions and choose from a menu of topics in a self-service manner by pressing the corresponding keys on their phones (eg, *"Thank you for calling Butabika Hospital. Please choose your preferred language. For English, press 1, Bw'oba oyagala kuwuliriza mu Luganda, nyiga 2"*). Figure 2 shows the IVR flow and the topics covered.

XSL•FO
RenderX

Figure 2. Mental health information topics and how they are accessed in the interactive voice response (IVR).



From the IVR, callers can choose to speak directly (live call) with an agent, for example, to seek more clarification on information in the IVR or ask for information that is not covered by the IVR. If it is during office hours, the system connects the caller to the agent. We had a total of 8 agents comprising 2 PSWs, 1 psychologist, 2 psychiatric clinical officers, and 3 psychiatric nurses. The staff do not sit in a physical call center; rather, they are accessible via dedicated mobile phones. All their phones are dialed concurrently (“ring all” strategy), and whoever picks first responds to the caller. Outside working hours, callers are instructed by the system to leave a voicemail, and the call agents returns the calls the next day; this is only possible from a softphone on a computer within the hospital since the caller’s number is hidden on the agents’ mobile phones for privacy reasons. All agents were encouraged to respond to the calls immediately, and a schedule was created for responding to voicemails. The psychologist provided supervision to the agents and handled any difficult cases, which the agents were encouraged to report or escalate whenever necessary.

Development of the system began by developing the IVR content (mental health messages and navigation instructions), which was done concurrently with the requirements specification process described in *Step 2: Specifying the Requirements of the System* section. The team iteratively wrote the script for mental health messages based on their clinical expertise, reviewed the Luganda translation, and discussed the IVR menu options and caller-system interaction based on the requirements and insights from prior work. We limited the IVR options to a manageable number and organized the information in a logical order, that is, from general information (overview of mental illnesses) to specific information (eg, individual illness such as anxiety or depression). Attention was paid to ease of language (eg, description of concepts or illnesses in addition to naming them and reduction in use of medical jargon); tone (calm, empathic, and nonjudgmental); and cultural appropriateness (eg, acknowledging the role of faith and alternative medicine). The developed content was recorded in a professional audio recording studio and deployed in private branch exchange (PBX) software by the IT professional and the first author.

Results

The telephone system was implemented using Issabel (Issabel LLC), an open-source PBX software based on Asterisk (Sangoma Technologies Corp). It was deployed on a simple server (Intel Core i5 2.6 GHz, 8 GB RAM, 1-TB Hard disk) at Butabika Hospital and connected to a local mobile telecom provider via session initiation protocol with 12 trunks. The calls to the system are reverse billed and therefore are free to the callers.

In sum, we developed a total of 22 messages, 14 (64%) of which were on mental health or other practical information needs elicited from the participants, that is, overview of what mental illnesses are; the causes, signs, and symptoms; the common mental illnesses in Uganda; assessment and management of mental illness; and how to navigate the health care system. The remaining 8 (36%) messages contained navigation instructions

XSL•FO
RenderX

or feedback to user (welcome message, language selection, disclaimer, warning in case of emergency, the different menu options, invalid selection, message replay, returning to main menu, and voicemail instructions). The messages were then translated to Luganda for a total of 44 messages. Figure 2 shows the topics addressed by the IVR messages (without some of the navigation messages).

Step 4: Demonstration and Evaluation

Procedure and Participants

Following deployment, we held a 1-day workshop with the PSWs, nurses, and psychiatric clinical officers (n=10) who had participated in the FGDs to test the system, get feedback about the IVR content, and identify and correct any system malfunctions or errors (eg, if there were language mix up or a wrong response for a particular IVR option chosen by the caller).

We held a second workshop to train the call agents on workflows, software system, and phone etiquettes and how to communicate with persons with mental health problems. We also discussed operational issues, for example, definition of office hours when live calls should be allowed, and schedules for returning voicemails and evaluation survey.

The system was advertised via the hospital website and social media channels, posters in the hospital, and personal contacts of the staff and participants. After go-live, we continued to supervise the call agents and held regular review meetings in which we listened to recorded calls and critiqued the conversations, offered support to the call agents (especially the PSW) in case of difficult calls, and collected feedback on usability and user perception of utility of the service.

Results

No major problems were found during the testing workshops, but participants reported that the workshops helped them better understand the service from practically trying it out. They showed enthusiasm for their roles as call agents and became ambassadors who advertised the service to patients and their social networks. Schedules were also drawn for returning calls and office hours defined, which were then programmed into the PBX, sending calls outside these hours to voicemail.

The system went live in August 2021. Detailed results from a survey of the callers and analysis of the use patterns will be reported in a separate study (under preparation), but here we summarize the observations from the first 4 months of operation.

From August to December 2022, a total of 456 calls, from 236 unique numbers (average of 4 calls per day), were made to the system, that is, reaching at least the IVR (automated) component. Of these, 99 (21.7%) calls were made during out-of-office hours for the call agents, so they went to voicemail and were called back within the following days. Of the remaining 357 calls made during office hours, 80 (22.4%) calls stopped at the IVR, while 231 (64.7%) proceeded to speak to a live agent (note that the percentages do not add up to 100% because some callers made multiple calls using the IVR or leaving a voicemail and later called and spoke to a live agent). Furthermore, the 22 (6.2%) calls were never answered by the call agents. On average, live calls were answered within 11 (SD 7) seconds, and their average length was 3.5 (SD 2.8) minutes.

Callers came from all parts of the country (as far as 8 hours by road from Butabika Hospital), although the majority were from the central region (within a 1-hour distance from Butabika Hospital). They included caregivers seeking advice about relatives who were showing symptoms of mental illnesses or those already undergoing care; mental health patients who were relatively stable and were seeking advice about medication or return dates; and others such as clinicians from other health facilities, journalists, and government officials who wanted more information about the call center system or the mental health care services offered at Butabika Hospital. Calls about patients who had “escaped” from the hospital were also common, often made by concerned community members near the hospital who come across a person with mental illness wandering in the community. Generally, the service has been received positively. Callers were especially happy with the PSWs who shared their personal journeys with mental illnesses and recovery, and this encouraged them to overcome the stigma and negativity that they had about mental health care services. The PSWs also reported positive experiences, stating that working as call agents and helping others gave them a sense of purpose and brought order and calmness.

A key challenge was callers who required specific and individualized information that the call agents did not have at hand and could not be prerecorded in the IVR. Such information included requests for prescriptions, questions on stocks of certain medications, availability and cost of certain tests and procedures, or about the condition of a relative who was admitted in the hospital.

Reflexivity

The members of the research team who were involved in data collection and analysis (FGDs, workshops, and analysis meetings) are intimately familiar with the local context and understand participants’ realities (including participants’ access and use of mobile phones and the internet) since they come from the same region of the country and speak the local language (Luganda). This made it easier to communicate with the participants (even for sessions that were held in Luganda) and to understand and relate to the ideas or issues they raised. To reduce potential undue coercion, the clinicians involved in the care of the participants (patients and PSWs) did not participate in the FGDs sessions but participated in data analysis and interpretation. These clinicians were especially important in ensuring that the rest of the research team members were aware of assumptions and potential prejudices, for example, with regard to beliefs in witchcraft as a cause for mental illnesses or in faith healing, common among those with low education status. Clinicians working in mental health care in this context frequently encounter such beliefs and appreciate the importance of respecting them, which was also useful for informing how we crafted the mental health messages in the system. Moreover, 3 of the research team members were from a different high-income country and brought in different perspectives, which helped us question our interpretations and assumptions.

XSL•FO
RenderX

Discussion

Principal Findings

This paper describes the development and implementation of a digital health intervention aimed at improving mental health care in Uganda. Using principles of UCD [18-20] and DSR [21,22], we systematically engaged stakeholders, collected data on target users' experiences of mental health care, their opinions and recommendations about the proposed mental health telephone service, and contextual issues that could influence implementation. We used the CFIR, an established implementation science meta-framework [26-31], to collect and analyze these data and derive system requirements and then iteratively cocreated and tested the system.

We identified several challenges faced by patients with mental health problems and their caregivers in Uganda and peculiarities about the organization and the wider societal context, which supported the proposed innovation. These challenges included the severe shortage of mental health workers and services, lack of awareness, negative beliefs and norms, stigma, huge burden of caregiving, and financial challenges. At the same time, there is a general trend toward digitalization of health care to improve patient experience and efficiency of health care, and participants were enthusiastic about our proposed call center because they were familiar with the technology and considered it as a feasible, affordable, convenient, and efficient way to get mental health services without being stigmatized. The participants also gave several recommendations on how to successfully implement the intervention, for example, by making calls toll free, ensuring 24×7 availability, providing mental health information in multiple languages, using technologies or channels that are appropriate to the context (telephone calls and IVR), sufficient staffing to reducing call waiting times, sensitizing people about the service, and training and supervision of the call agents to ensure quality service. Early evaluation of the intervention shows that clients are very positive about the service, particularly with the use of PSWs (recovering patients) who share their lived experience with others.

Comparison With Prior Work

Prior research has demonstrated the value of mobile health (mHealth) in addressing some of the health care challenges in Uganda and similar contexts elsewhere. Systematic reviews on mHealth in general [39-45] or on specific clinical domains such as HIV or AIDS [45-47] and palliative care [48] have highlighted the improvement of health care coordination and communication between patients and health care providers, patient adherence to treatment and reduction of loss to follow-up, patient engagement and self-care, facilitation of community-based care, and improvement of access to care for rural or geographically isolated populations. Advantages such as ubiquity of mobile technology, affordability and acceptability by patients and health workers, interactivity and personalization, and saving of time and cost of traveling to health facilities have been cited. Examples of prior studies on mHealth in Uganda include use of IVR, SMS text messages, and phone calls to support the management of HIV or AIDS [49,50] and tuberculosis [51]; use of IVR to address barriers to fistula care in Uganda [52]; SMS text messages for stroke rehabilitation [53]; and IVR for provision of cancer awareness and advice [38]. There is also a commercial digital health company that has operated different mHealth services in Uganda for approximately 10 years [54]. Unfortunately, the use of mHealth in mental health in Uganda and Africa in general is limited [8,55,56]. This is likely due to the general underfunding of mental health care services [10-14]. Available research on mHealth in mental health is mostly from developed countries [57-61], with many interventions using the internet and smartphone apps, which might not be accessible or affordable in Uganda or other LMICs. Interventions that use basic phone features such as SMS text messages, IVR, and voice calls are more appropriate in LMICs as they overcome infrastructural limitations. Such interventions are also relevant for low-income and migrant communities in developed countries since these populations face low digital health literacy and language barriers [62-65].

In the previous project led by the first author for the provision of cancer information [38], similar findings in terms of challenges faced by patients, requirements and recommendations for the system, and generally positive reception after implementation were reported. The cancer awareness system mainly used the IVR feature with prerecorded information, with the option to speak to a live agent added as an emergency due to the COVID-19 pandemic. The agents were health care workers (nurses and physicians) who, due to travel restrictions, had been free to handle phone calls. While callers appreciated this feature, it is otherwise not possible given the limited number of health workers. In this study, PSWs helped to address the shortage of health care professionals. A large multinational research study from Uganda and elsewhere has demonstrated the positive benefits of using PSWs, both for their own recovery and for the health care system [66-68]. Our study builds onto this prior work to innovatively and efficiently put this underused resource to use through digital health.

Strengths and Limitations of the Study

A strength of this study is the strong theoretical underpinning. Implementation studies have been faulted in the past for not being theory driven, which undermines the adoption of digital technologies [30]. The UCD and DSR approach used informed a systematic cocreation process of intervention development with user participation, while the CFIR allowed a comprehensive review of user, technological, and contextual issues to inform system requirements. Even so, we could not consider all the requirements or recommendations by the participants when designing the system because of resource limitations or contradictions. For example, some participants recommended adding video calling features to the system to enhance interaction and assessment of affect. Other participants had concerns about continuity of care, which indeed is difficult to achieve with the current call center system that lacks electronic medical records or mechanisms to ensure that callers are directed to agents with whom they have interacted with before. However, adding such features would make the system complex, expensive, and inaccessible to some such as those who mentioned inability to work with smartphones or had connectivity problems. Still, the insights from this

XSL•FO
RenderX

comprehensive assessment can inform future incremental iterations of the system during scale-up.

Conclusions

Participants were enthusiastic about the proposed call center because they were familiar with the technology and considered it as a feasible, affordable, convenient, and efficient way to get mental health services without being stigmatized. The system provided mental health information and linkage to health care providers and PSWs. The information in audio format made it accessible even to the people with low literacy, and the automated IVR allowed 24×7 access while reducing the pressure on the health care workforce. Translation to English and Luganda, the 2 most spoken languages in Uganda, increased reach, as did the reverse billing (no cost to the caller) and the use of basic telephone calls as the channel of access since many Ugandans still do not have affordable and reliable internet access.

Recommendations

In this study, people with mental illness, caregivers, and health care providers deemed a telephone-based mental health care service useful and necessary to increase access to mental health information and care and reduce stigma toward people with mental health problems. This positive view needs to be harnessed to scale up the digitalization of mental health care including providing therapy and establishing it in other mental health care settings in line with the current Ugandan digitalization policy and the Third National Development Plan. This method of mental health care may be replicable and scalable in other LMICs with mental health care system and personnel challenges similar to Uganda. Further research is needed to evaluate long-term adoption, patterns of use, and impact on clinical outcomes.

Acknowledgments

This project was funded by Digital Futures, a research collaboration between KTH Royal Institute of Technology, Stockholm University, and RISE Research Institutes of Sweden. Funders had no influence in the conduct of the research study.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Summary of qualitative findings coded according to the Consolidated Framework for Implementation Research.
[DOCX File , 17 KB - humanfactors_v11i1e53976_app1.docx]

Multimedia Appendix 2
Consolidated Framework for Implementation Research domains, constructs, themes from focus group discussions, and illustrative quotes.
[XLSX File (Microsoft Excel File), 27 KB - humanfactors_v11i1e53976_app2.xlsx]

Multimedia Appendix 3
Requirements and how they are addressed by the system and its setup.
[DOCX File , 18 KB - humanfactors_v11i1e53976_app3.docx]

References

1. Ingrid D, Brunier A, Sharma H. World mental health day: an opportunity to kick-start a massive scale-up in investment in mental health. World Health Organization. 2020. URL: https://tinyurl.com/2z6bramw [accessed 2024-04-29]
2. Kovacevic R. Mental health: lessons learned in 2020 for 2021 and forward. World Bank Blog. 2021. URL: https://blogs.worldbank.org/health/mental-health-lessons-learned-2020-2021-and-forward [accessed 2024-04-29]
3. Naslund JA, Aschbrenner KA, Araya R, Marsch LA, Unützer J, Patel V, et al. Digital technology for treating and preventing mental disorders in low-income and middle-income countries: a narrative review of the literature. Lancet Psychiatry 2017 Jun;4(6):486-500. [doi: 10.1016/s2215-0366(17)30096-2]
4. GBD 2019 Mental Disorders Collaborators. Global, regional, and national burden of 12 mental disorders in 204 countries and territories, 1990–2019: a systematic analysis for the Global Burden of Disease Study 2019. Lancet Psychiatry 2022 Feb;9(2):137-150. [doi: 10.1016/s2215-0366(21)00395-3]
5. The WHO special initiative for mental health (2019-2023): universal health coverage for mental health. World Health Organization. 2019. URL: https://apps.who.int/iris/handle/10665/310981 [accessed 2024-04-29]
6. Mental health. World Health Organization. URL: https://www.who.int/health-topics/mental-health#tab=tab_1 [accessed 2024-04-29]

XSL•FO
RenderX

7. Mental health and the 2030 sustainable development agenda: global inaction on mental health is putting the brakes on development. Mental Health Innovation Network. 2018. URL: https://www.mhinnovation.net/sites/default/files/content/document/Policy%20brief%202030%20SDG.pdf [accessed 2024-04-29]
8. Ben-Zeev D. Mobile health for mental health in west Africa: the case for Ghana. Psychiatr Serv 2018 Jul 01;69(7):741-743. [doi: 10.1176/appi.ps.201700555] [Medline: 29540120]
9. Arias D, Saxena S, Verguet S. Quantifying the global burden of mental disorders and their economic value. EClinicalMedicine 2022 Dec;54:101675 [FREE Full text] [doi: 10.1016/j.eclinm.2022.101675] [Medline: 36193171]
10. Kagaari J. Mental health in Uganda. American Psychological Association. 2021. URL: https://www.apa.org/international/global-insights/uganda-mental-health [accessed 2024-04-29]
11. Molodynski A, Cusack C, Nixon J. Mental healthcare in Uganda: desperate challenges but real opportunities. BJPsych Int 2017 Nov 02;14(4):98-100 [FREE Full text] [doi: 10.1192/s2056474000002129] [Medline: 29093962]
12. Alloh FT, Regmi P, Onche I, van Teijlingen E, Trenoweth S. Mental Health in low-and middle income countries (LMICs): going beyond the need for funding. Health Prospect 2018 Jun 19;17(1):12-17. [doi: 10.3126/hprospect.v17i1.20351]
13. Gberie L. Mental illness: invisible but devastating. Africa Renewal. URL: https://tinyurl.com/mwhvxjzu [accessed 2024-04-29]
14. Patel V, Maj M, Flisher AJ, De Silva MJ, Koschorke M, Prince M, WPA ZonalMember Society Representatives. Reducing the treatment gap for mental disorders: a WPA survey. World Psychiatry 2010 Oct;9(3):169-176 [FREE Full text] [doi: 10.1002/j.2051-5545.2010.tb00305.x] [Medline: 20975864]
15. Ezeugwu CR, Ojedokun O. Masculine norms and mental health of African men: what can psychology do? Heliyon 2020 Dec;6(12):e05650. [doi: 10.1016/j.heliyon.2020.e05650]
16. Esan O, Abdumalik J, Eaton J, Kola L, Fadahunsi W, Gureje O. Mental health care in Anglophone West Africa. Psychiatr Serv 2014 Sep 01;65(9):1084-1087. [doi: 10.1176/appi.ps.201300300] [Medline: 25179185]
17. Musisi S, Kinyanda E. Long-term impact of war, civil war, and persecution in civilian populations-conflict and post-traumatic stress in African communities. Front Psychiatry 2020 Feb 25;11:20 [FREE Full text] [doi: 10.3389/fpsyt.2020.00020] [Medline: 32158407]
18. Abras C, Maloney-Krichmar D, Preece J. User-centered design. In: Bainbridge W, editor. Encyclopedia of Human-Computer Interaction. Thousand Oaks, CA: SAGE Publications; 2004:445-456.
19. De Vito Dabbs A, Myers BA, Mc Curry KR, Dunbar-Jacob J, Hawkins RP, Begey A, et al. User-centered design and interactive health technologies for patients. Comput Inform Nurs 2009;27(3):175-183 [FREE Full text] [doi: 10.1097/NCN.0b013e31819f7c7c] [Medline: 19411947]
20. Mao JY, Vredenburg K, Smith PW, Carey T. The state of user-centered design practice. Commun ACM 2005 Mar;48(3):105-109. [doi: 10.1145/1047671.1047677]
21. Peffers K, Tuunanen T, Rothenberger MA, Chatterjee S. A design science research methodology for information systems research. J Manag Inf Syst 2014 Dec 08;24(3):45-77. [doi: 10.2753/mis0742-1222240302]
22. Iivari J, Venable JR. Action research and design science research - seemingly similar but decisively dissimilar. In: Proceedings of the 17th European Conference on Information Systems. 2009 Presented at: ECIS '09; June 8-10, 2009; Verona, Italy URL: https://aisel.aisnet.org/cgi/viewcontent.cgi?article=1025&context=ecis2009
23. Brooke J. SUS: a 'quick and dirty' usability scale. In: Jordan PW, Thomas B, McClelland IL, Weerdmeester B, editors. Usability Evaluation In Industry. Boca Raton, FL: CRC Press; 1996:207-212.
24. Brooke J. SUS: a retrospective. J Usability Stud 2013:29-40 [FREE Full text]
25. Smart A. A multi-dimensional model of clinical utility. Int J Qual Health Care 2006 Oct;18(5):377-382. [doi: 10.1093/intqhc/mzl034] [Medline: 16951425]
26. Damschroder LJ, Aron DC, Keith RE, Kirsh SR, Alexander JA, Lowery JC. Fostering implementation of health services research findings into practice: a consolidated framework for advancing implementation science. Implement Sci 2009 Aug 07;4(1):50 [FREE Full text] [doi: 10.1186/1748-5908-4-50] [Medline: 19664226]
27. Kirk MA, Kelley C, Yankey N, Birken SA, Abadie B, Damschroder L. A systematic review of the use of the consolidated framework for implementation research. Implement Sci 2016 May 17;11(1):72 [FREE Full text] [doi: 10.1186/s13012-016-0437-z] [Medline: 27189233]
28. Skolarus TA, Lehmann T, Tabak RG, Harris J, Lecy J, Sales AE. Assessing citation networks for dissemination and implementation research frameworks. Implement Sci 2017 Jul 28;12(1):97 [FREE Full text] [doi: 10.1186/s13012-017-0628-2] [Medline: 28754140]
29. Damschroder LJ, Reardon CM, Widerquist MA, Lowery J. The updated consolidated framework for implementation research based on user feedback. Implement Sci 2022 Oct 29;17(1):75 [FREE Full text] [doi: 10.1186/s13012-022-01245-0] [Medline: 36309746]
30. Nilsen P. Making sense of implementation theories, models and frameworks. Implement Sci 2015 Apr 21;10(1):53 [FREE Full text] [doi: 10.1186/s13012-015-0242-0] [Medline: 25895742]
31. Damschroder LJ. Clarity out of chaos: use of theory in implementation research. Psychiatry Res 2020 Jan;283:112461 [FREE Full text] [doi: 10.1016/j.psychres.2019.06.036] [Medline: 31257020]

XSL•FO
RenderX

32. Damschroder LJ, Reardon CM, Opra Widerquist MA, Lowery J. Conceptualizing outcomes for use with the Consolidated Framework for Implementation Research (CFIR): the CFIR outcomes addendum. Implement Sci 2022 Jan 22;17(1):7 [FREE Full text] [doi: 10.1186/s13012-021-01181-5] [Medline: 35065675]
33. Hsieh HF, Shannon SE. Three approaches to qualitative content analysis. Qual Health Res 2005 Nov 01;15(9):1277-1288. [doi: 10.1177/1049732305276687] [Medline: 16204405]
34. Suchman L, Omoluabi E, Kramer J, Vallin J, Sedlander E, Gitome S, ICAN Research Consortium. Analyzing fast and slow: combining traditional and rapid qualitative analysis to meet multiple objectives of a complex transnational study. Front Sociol 2023 Feb 1;8:961202 [FREE Full text] [doi: 10.3389/fsoc.2023.961202] [Medline: 36818663]
35. Lewinski AA, Crowley MJ, Miller C, Bosworth HB, Jackson GL, Steinhauser K, et al. Applied rapid qualitative analysis to develop a contextually appropriate intervention and increase the likelihood of uptake. Med Care 2021 Jun 01;59(Suppl 3):S242-S251 [FREE Full text] [doi: 10.1097/MLR.0000000000001553] [Medline: 33976073]
36. Gale RC, Wu J, Erhardt T, Bounthavong M, Reardon CM, Damschroder LJ, et al. Comparison of rapid vs in-depth qualitative analytic methods from a process evaluation of academic detailing in the Veterans Health Administration. Implement Sci 2019 Feb 01;14(1):11 [FREE Full text] [doi: 10.1186/s13012-019-0853-y] [Medline: 30709368]
37. Buetow S. Thematic analysis and its reconceptualization as 'saliency analysis'. J Health Serv Res Policy 2010 Apr 01;15(2):123-125. [doi: 10.1258/jhsrp.2009.009081] [Medline: 19762883]
38. Kabukye JK, Ilozumba O, Broerse JE, de Keizer N, Cornet R. Implementation of an interactive voice response system for cancer awareness in Uganda: mixed methods study. JMIR Mhealth Uhealth 2021 Jan 26;9(1):e22061 [FREE Full text] [doi: 10.2196/22061] [Medline: 33496672]
39. McCool J, Dobson R, Whittaker R, Paton C. Mobile health (mHealth) in low- and middle-income countries. Annu Rev Public Health 2022 Apr 05;43(1):525-539 [FREE Full text] [doi: 10.1146/annurev-publhealth-052620-093850] [Medline: 34648368]
40. Hurt K, Walker RJ, Campbell JA, Egede LE. mHealth interventions in low and middle-income countries: a systematic review. Glob J Health Sci 2016 Sep 01;8(9):54429 [FREE Full text] [doi: 10.5539/gjhs.v8n9p183] [Medline: 27157176]
41. Chib A, van Velthoven MH, Car J. mHealth adoption in low-resource environments: a review of the use of mobile healthcare in developing countries. J Health Commun 2015 Mar 27;20(1):4-34. [doi: 10.1080/10810730.2013.864735] [Medline: 24673171]
42. Aranda-Jan CB, Mohutsiwa-Dibe N, Loukanova S. Systematic review on what works, what does not work and why of implementation of mobile health (mHealth) projects in Africa. BMC Public Health 2014 Feb 21;14(1):188 [FREE Full text] [doi: 10.1186/1471-2458-14-188] [Medline: 24555733]
43. Hall CS, Fottrell E, Wilkinson S, Byass P. Assessing the impact of mHealth interventions in low- and middle-income countries--what has been shown to work? Glob Health Action 2014;7:25606 [FREE Full text] [doi: 10.3402/gha.v7.25606] [Medline: 25361730]
44. Abreu FD, Bissaco MA, Silva AP, Boschi SR, Scardovelli TA, Santos MF, et al. The use and impact of mHealth by community health workers in developing and least developed countries: a systematic review. Res Biomed Eng 2021 May 20;37(3):563-582 [FREE Full text] [doi: 10.1007/S42600-021-00154-3]
45. Cooper V, Clatworthy J, Whetham J, Consortium E. mHealth interventions to support self-management in HIV: a systematic review. Open AIDS J 2017 Nov 21;11(1):119-132 [FREE Full text] [doi: 10.2174/1874613601711010119] [Medline: 29290888]
46. Devi BR, Syed-Abdul S, Kumar A, Iqbal U, Nguyen PA, Li YC, et al. mHealth: an updated systematic review with a focus on HIV/AIDS and tuberculosis long term management using mobile phones. Comput Methods Programs Biomed 2015 Nov;122(2):257-265 [FREE Full text] [doi: 10.1016/j.cmpb.2015.08.003] [Medline: 26304621]
47. Catalani C, Philbrick W, Fraser H, Mechael P, Israelski DM. mHealth for HIV treatment and prevention: a systematic review of the literature. Open AIDS J 2013 Sep 20;7(1):17-41 [FREE Full text] [doi: 10.2174/1874613620130812003] [Medline: 24133558]
48. Allsop MJ, Powell RA, Namisango E. The state of mHealth development and use by palliative care services in sub-Saharan Africa: a systematic review of the literature. BMJ Support Palliat Care 2018 Jun 20;8(2):155-163 [FREE Full text] [doi: 10.1136/bmjspcare-2015-001034] [Medline: 27207721]
49. Byonanebye DM, Nabaggala MS, Naggirinya AB, Lamorde M, Oseku E, King R, et al. An interactive voice response software to improve the quality of life of people living with HIV in Uganda: randomized controlled trial. JMIR Mhealth Uhealth 2021 Feb 11;9(2):e22229 [FREE Full text] [doi: 10.2196/22229] [Medline: 33570497]
50. Twimukye A, Bwanika Naggirinya A, Parkes-Ratanshi R, Kasirye R, Kiragga A, Castelnuovo B, et al. Acceptability of a mobile phone support tool (call for life Uganda) for promoting adherence to antiretroviral therapy among young adults in a randomized controlled trial: exploratory qualitative study. JMIR Mhealth Uhealth 2021 Jun 14;9(6):e17418 [FREE Full text] [doi: 10.2196/17418] [Medline: 34121665]
51. Byonanebye DM, Mackline H, Sekaggya-Wiltshire C, Kiragga AN, Lamorde M, Oseku E, et al. Impact of a mobile phone-based interactive voice response software on tuberculosis treatment outcomes in Uganda (CFL-TB): a protocol for a randomized controlled trial. Trials 2021 Jun 13;22(1):391 [FREE Full text] [doi: 10.1186/s13063-021-05352-z] [Medline: 34120649]

XSL•FO
RenderX

52. Tripathi V, Arnoff E, Bellows B, Sripad P. Use of interactive voice response technology to address barriers to fistula care in Nigeria and Uganda. Mhealth 2020 Apr;6:12 [FREE Full text] [doi: 10.21037/mhealth.2019.12.04] [Medline: 32270004]
53. Teriö M, Eriksson G, Kamwesiga JT, Guidetti S. What's in it for me? A process evaluation of the implementation of a mobile phone-supported intervention after stroke in Uganda. BMC Public Health 2019 May 14;19(1):562. [doi: 10.1186/s12889-019-6849-3] [Medline: 31088411]
54. Kamulegey L, Ssebwana J, Abigaba W, Bwanika J, Musinguzi D. Mobile health in Uganda: a case study of the medical concierge group. East Afr Sci 2019 Mar 25;1(1):9-14. [doi: 10.24248/easci.v1i1.12]
55. Pearson AL, Mack E, Namanya J. Mobile phones and mental well-being: initial evidence suggesting the importance of staying connected to family in rural, remote communities in Uganda. PLoS One 2017 Jan 17;12(1):e0169819 [FREE Full text] [doi: 10.1371/journal.pone.0169819] [Medline: 28095427]
56. Adepoju P. Africa turns to telemedicine to close mental health gap. Lancet Digit Health 2020 Nov;2(11):e571-e572. [doi: 10.1016/s2589-7500(20)30252-1]
57. Lattie EG, Stiles-Shields C, Graham AK. An overview of and recommendations for more accessible digital mental health services. Nat Rev Psychol 2022 Feb 26;1(2):87-100 [FREE Full text] [doi: 10.1038/s44159-021-00003-1] [Medline: 38515434]
58. Price M, Yuen EK, Goetter EM, Herbert JD, Forman EM, Acierno R, et al. mHealth: a mechanism to deliver more accessible, more effective mental health care. Clin Psychol Psychother 2014 Aug 05;21(5):427-436 [FREE Full text] [doi: 10.1002/cpp.1855] [Medline: 23918764]
59. Wang K, Varma DS, Prosperi M. A systematic review of the effectiveness of mobile apps for monitoring and management of mental health symptoms or disorders. J Psychiatr Res 2018 Dec;107:73-78 [FREE Full text] [doi: 10.1016/j.jpsychires.2018.10.006] [Medline: 30347316]
60. Hilty DM, Ferrer DC, Parish MB, Johnston B, Callahan EJ, Yellowlees PM. The effectiveness of telemental health: a 2013 review. Telemed J E Health 2013 Jun;19(6):444-454 [FREE Full text] [doi: 10.1089/tmj.2013.0075] [Medline: 23697504]
61. Hwang WJ, Ha JS, Kim MJ. Research trends on mobile mental health application for general population: a scoping review. Int J Environ Res Public Health 2021 Mar 02;18(5):2459 [FREE Full text] [doi: 10.3390/ijerph18052459] [Medline: 33801537]
62. Woodward A, Burchert S, Barry AS, Broerse JE, Sondorp E, Bold A, et al. Scalability of digital psychological innovations for refugees: a comparative analysis in Egypt, Germany, and Sweden. SSM Mental Health 2023 Dec;4:100231. [doi: 10.1016/j.ssmmh.2023.100231]
63. Bergman L, Nilsson U, Dahlberg K, Jaensson M, Wångdahl J. Health literacy and e-health literacy among Arabic-speaking migrants in Sweden: a cross-sectional study. BMC Public Health 2021 Nov 25;21(1):2165 [FREE Full text] [doi: 10.1186/s12889-021-12187-5] [Medline: 34823499]
64. Reuland CJ, Godage SK, Wu L, Valenzuela-Araujo D, Cortez JD, Polk S, et al. Information and communication technology access and use among low-income Latino immigrant parents. Matern Child Health J 2021 Dec 23;25(12):1807-1813. [doi: 10.1007/s10995-021-03265-6] [Medline: 34687401]
65. Bender MS, Choi J, Arai S, Paul SM, Gonzalez P, Fukuoka Y. Digital technology ownership, usage, and factors predicting downloading health apps among Caucasian, Filipino, Korean, and Latino Americans: the digital link to health survey. JMIR Mhealth Uhealth 2014 Oct 22;2(4):e43 [FREE Full text] [doi: 10.2196/mhealth.3710] [Medline: 25339246]
66. Puschner B, Repper J, Mahlke C, Nixdorf R, Basangwa D, Nakku J, et al. Using peer support in developing empowering mental health services (UPSIDES): background, rationale and methodology. Ann Glob Health 2019 Apr 05;85(1):53 [FREE Full text] [doi: 10.5334/aogh.2435] [Medline: 30951270]
67. Moran GS, Kalha J, Mueller-Stierlin AS, Kilian R, Krumm S, Slade M, et al. Peer support for people with severe mental illness versus usual care in high-, middle- and low-income countries: study protocol for a pragmatic, multicentre, randomised controlled trial (UPSIDES-RCT). Trials 2020 May 01;21(1):371 [FREE Full text] [doi: 10.1186/s13063-020-4177-7] [Medline: 32357903]
68. Nixdorf R, Nugent L, Aslam R, Barber S, Charles A, Gai Meir L, et al. A multi-national peer support intervention: the UPSIDES pilot phase. Adv Ment Health 2022 Jan 06;20(1):2-14 [FREE Full text] [doi: 10.1080/18387357.2021.2020140]

Abbreviations

CFIR: Consolidated Framework for Implementation Research
DSR: design science research
FGD: focus group discussion
IVR: interactive voice response
LMIC: low- and middle-income country
mHealth: mobile health
PBX: private branch exchange
PSW: peer support worker
UCD: user-centered design

XSL•FO
RenderX

Edited by A Choudhury; submitted 25.10.23; peer-reviewed by A Gangadhara Rao, I Wilson; comments to author 23.02.24; revised version received 01.04.24; accepted 01.05.24; published 06.06.24.

Please cite as:
Kabukye JK, Namagembe R, Nakku J, Kiberu V, Sjölinder M, Nilsson S, Wamala-Larsson C
Implementing a Hospital Call Center Service for Mental Health in Uganda: User-Centered Design Approach
JMIR Hum Factors 2024;11:e53976
URL: https://humanfactors.jmir.org/2024/1/e53976
doi:10.2196/53976
PMID:38843515

©Johnblack K Kabukye, Rosemary Namagembe, Juliet Nakku, Vincent Kiberu, Marie Sjölinder, Susanne Nilsson, Caroline Wamala-Larsson. Originally published in JMIR Human Factors (https://humanfactors.jmir.org), 06.06.2024. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on https://humanfactors.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Evaluation of a Computer-Aided Clinical Decision Support System for Point-of-Care Use in Low-Resource Primary Care Settings: Acceptability Evaluation Study

Geletaw Sahle Tegenaw[1,2], PhD; Demisew Amenu Sori[3], MD; Girum Ketema Teklemariam[2], PhD; Frank Verbeke[1], MD, PhD; Jan Cornelis[1], PhD, Prof Dr; Bart Jansen[1,4], PhD, Prof Dr

[1]Department of Electronics and Informatics, Vrije Universiteit Brussel, Brussel, Belgium
[2]Faculty of Computing, Jimma Institute of Technology, Jimma University, Jimma, Ethiopia
[3]Department of Obstetrics and Gynecology, College of Health Science, Jimma University, Jimma, Ethiopia
[4]Interuniversitair Micro-Electronica Centrum, Leuven, Belgium

Corresponding Author:
Geletaw Sahle Tegenaw, PhD
Department of Electronics and Informatics
Vrije Universiteit Brussel
Pleinlaan 2
Brussel, 1050
Belgium
Phone: 32 485744961
Email: gtegenaw@vub.be

Abstract

Background: A clinical decision support system (CDSS) based on the logic and philosophy of clinical pathways is critical for managing the quality of health care and for standardizing care processes. Using such a system at a point-of-care setting is becoming more frequent these days. However, in a low-resource setting (LRS), such systems are frequently overlooked.

Objective: The purpose of the study was to evaluate the user acceptance of a CDSS in LRSs.

Methods: The CDSS evaluation was carried out at the Jimma Health Center and the Jimma Higher Two Health Center, Jimma, Ethiopia. The evaluation was based on 22 parameters organized into 6 categories: ease of use, system quality, information quality, decision changes, process changes, and user acceptance. A Mann-Whitney U test was used to investigate whether the difference between the 2 health centers was significant (2-tailed, 95% CI; α=.05). Pearson correlation and partial least squares structural equation modeling (PLS-SEM) was used to identify the relationship and factors influencing the overall acceptance of the CDSS in an LRS.

Results: On the basis of 116 antenatal care, pregnant patient care, and postnatal care cases, 73 CDSS evaluation responses were recorded. We found that the 2 health centers did not differ significantly on 16 evaluation parameters. We did, however, detect a statistically significant difference in 6 parameters (P<.05). PLS-SEM results showed that the coefficient of determination, R^2, of perceived user acceptance was 0.703. More precisely, the perceived ease of use (β=.015, P=.91) and information quality (β=.149, P=.25) had no positive effect on CDSS acceptance but, rather, on the system quality and perceived benefits of the CDSS, with P<.05 and β=.321 and β=.486, respectively. Furthermore, the perceived ease of use was influenced by information quality and system quality, with an R^2 value of 0.479, indicating that the influence of information quality on the ease of use is significant but the influence of system quality on the ease of use is not, with β=.678 (P<.05) and β=.021(P=.89), respectively. Moreover, the influence of decision changes (β=.374, P<.05) and process changes (β=.749, P<.05) both was significant on perceived benefits (R^2=0.983).

Conclusions: This study concludes that users are more likely to accept and use a CDSS at the point of care when it is easy to grasp the perceived benefits and system quality in terms of health care professionals' needs. We believe that the CDSS acceptance model developed in this study reveals specific factors and variables that constitute a step toward the effective adoption and deployment of a CDSS in LRSs.

(JMIR Hum Factors 2024;11:e47631) doi:10.2196/47631

XSL•FO
RenderX

KEYWORDS

low-resource setting; clinical decision support system; point-of-care instrument; evaluation; user acceptance; structural equation modeling; partial least squares structural equation modeling; decision-making; decision making; decision support; caregiver; users; acceptance; system quality

Introduction

The use of health information systems has considerably transformed the health care sector in recent years [1]. Proper and coordinated implementation is beneficial to the enhancement of health care delivery [2,3]. An effective clinical decision support system (CDSS); low-cost, point-of-care diagnostics; effective remote clinics; home-based therapies; and improved communication with patients and across health care facilities are among the benefits [4,5]. Even though the implementation of a CDSS at the point of care has sought to improve treatment quality and resource efficiency, its use in low-resource settings (LRSs) has lagged behind due to a variety of restrictions.

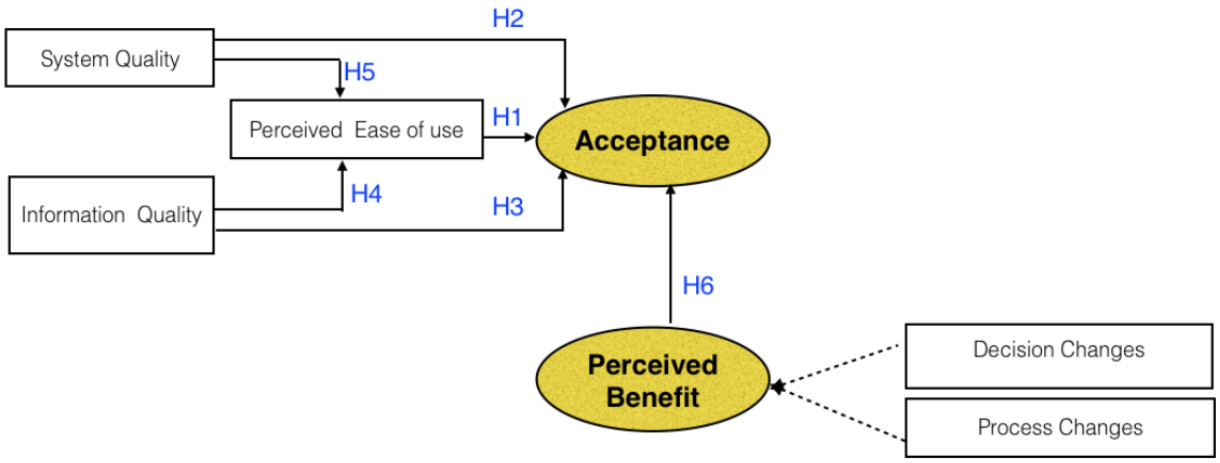
In Ethiopia, the health care system is a 3-tiered system organized into primary, secondary, and tertiary levels of care [6]. Primary health care settings include primary hospitals, health centers, and health posts. Recently, an electronic community health information system and district health information software were implemented in Ethiopian public health centers. These tools are commonly used for routine data management tasks. Frontline workers, however, lacked easy access to decision support systems and other similar point-of-care technologies. Paper-based clinical guidelines (CGs), card sheets, and point-of-care charts were the only available resources, and only limited information is documented on the card sheets [7,8]. Delivering evidence-based services at the point of care by capturing the required clinical data, summarizing and processing them in a consistent manner, and constructing a patient flow sheet to monitor and record the progress of care from the paper-based resources were challenging [7,8]. The Ethiopian national maturity health information assessment survey also revealed that there is a lack of health information infrastructure, a lack of decision support and knowledge management systems, and a lack of parameters and metrics for analyzing the impact of data [9].

Thus, introducing and integrating a CDSS with the existing health information system helps deliver appropriate, consistent, and integrated care. To introduce a CDSS in LRSs, we followed a 3-step approach:

- Step 1: A case study (maternal and childcare health services) needs analysis was conducted in LRSs to assess the available point-of-care evidence of the requirements for a CDSS, such as clinical pathways (CPs) or workflows [7,8].
- Step 2: We conducted a state-of-the-art review to investigate strategies and approaches for designing CDSS instruments for LRSs [10]. The aim was to review existing publications in the LRS context to explore recommended approaches and design considerations for building a CDSS.
- Step 3: A CDSS was developed based on the findings of the needs analysis and a review of the state of the art. The CDSS was designed to reduce delays and support frontline workers. The proposed CP algorithm, in particular, aims to find referrals and locally treatable cases by integrating knowledge-based approaches and historical evidence [11].

The aim of this study was to evaluate the user acceptance of a CDSS in LRSs. Overall, as depicted in Figure 1, this study proposed the following hypotheses to evaluate the user acceptance of the CDSS:

Figure 1. Computer-aided CDSS evaluation model hypotheses. Customized and adopted from Ji et al [20]. CDSS: Clinical decision support system; H: Hypothesis.



- Hypothesis (H)1: The perceived ease of use has a positive effect on the acceptance of a CDSS in LRSs.
- H2: System quality has a positive effect on the acceptance of a CDSS in LRSs.
- H3: Information quality has a positive effect on the acceptance of a CDSS in LRSs.
- H4: Information quality has a positive effect on the perceived ease of use of a CDSS in LRSs.

XSL•FO
RenderX

- H5: System quality has a positive effect on the perceived ease of use of a CDSS in LRSs.
- H6: Perceived benefits have a positive effect on the acceptance of CDSS in LRSs.

Methods

Ethical Considerations

Approval for the research was granted by the Institutional Review Board of the Institute of Health, Jimma University (reference number IHRPGI/467/19).

Study Settings and Participants

This study was conducted in low-resource primary health care centers, with a specific focus on the maternal and childcare health service units at the Jimma Health Center and the Jimma Higher Two Health Center. Both health centers are situated in Jimma Town, in the Oromia region, Southwestern Ethiopia. Each of them serves up to 40,000 people in its geographical area, accepts referrals from community health posts, and refers patients to the nearest hospital, such as the Shanan Gibe General Hospital and the Jimma University Specialized Hospital. The health centers serve and oversee both inpatient and outpatient cases. The number of personnel in the Jimma Higher Two Health Center is 34 and in the Jimma Health Center is 40, whereas in Ethiopia, the health center's maternal and child health service unit employs a much smaller number of health professionals, commonly 5-7 nurses and midwives. There were 5 nurses and midwives at the Jimma Health Center and 4 at the Jimma Higher Two Health Center during our investigation. The maternal and childcare health service unit is expected to serve 2000-2500 antenatal care (ANC), pregnant patient care, and postnatal care (PNC) cases annually.

Participants in the CDSS evaluation were health care professionals, such as midwives and nurses, who worked at the maternal and childcare health service unit at the Jimma Health Center and the Jimma Higher Two Health Center. The inclusion and exclusion criteria were as follows:

- Health care professionals were personnel at the maternal and childcare health service unit and were familiar with the existing clinical workflow, as well as volunteering to evaluate the CDSS.
- The ANC, pregnant patient care, and PNC cases that had been pre-recorded on the evaluation day were suitable for retrospective chart review to evaluate the CDSS.
- Both morning and afternoon evaluations were based on the pre-recorded cases from the respective morning and afternoon visits.

The CDSS evaluation was conducted in the health care professionals' spare time because the number of health care professionals at the maternal and childcare health service unit was limited, and they were so preoccupied and busy with their regular daily activities that it was not feasible to incorporate the evaluation into their routine. The health care professionals completed a questionnaire over the course of a half-day (as a summary of the half-day cases rather than as a case-by-case response), with the morning session taking place from 11:00 to1:00 A.M. and the afternoon session taking place from 5.00 to 18:30 P.M.

The initial evaluation was conducted in August 2022 at the Jimma Health Center. The second round of evaluation took place at the Jimma Health Center and the Jimma Higher Two Health Center from December 20, 2022, until January 15, 2023.

Based on our previous experience [13], obtaining the expected sample size in an LRS was difficult due to a shortage of health care professionals in the maternal and childcare health service unit (usually 4-7).

To determine the optimal strategies, we consulted the existing literature in support of our evaluation study design. Based on the findings of Mburu and Oboko's study [14], we observed that 79 cases were sufficient to assess the use of mobile health (mHealth) interventions in Kenya. Additionally, Mburu and Oboko [14] also reported that 60 subjects were sufficient to detect the small and medium effects of an exogenous latent variable (independent variable) on an endogenous latent variable (dependent variable), according to the findings of Chin and Newsted [15] and Cohen [16], just as using 40 subjects was sufficient for Goodhue et al [17]. The minimum sample size needed to observe an effect with a given power (ie, the probability of observing a statistically significant result at level P if a true effect of a certain magnitude is present) is determined by the effect size. The effect size is associated with the path coefficient between a variable that is assumed to describe a cause and a variable that is assumed to be an effect: values<0.02 indicate no effect, values>0.15 indicate a medium effect, and values>0.35 indicate a large effect [17,18]. Moreover, using 70-80 samples was adequate to model functional brain relationship hypotheses in the study by Sideridis et al [19]. However, Sideridis et al [19] also explicitly noted that sample sizes of 50 participants were associated with a root mean square error of approximation of <0.05, suggesting a satisfactory fit.

The study entailed a proof-of-principle CDSS evaluation using a convenience sample of 7 health professionals. Altogether, we reviewed 73 ANC, pregnant patient, and PNC cases.

Procedure and Measurement Instrument

A tutorial and a demonstration were provided to the health care professionals at the 2 health centers prior to using the CDSS. The health care professionals used and assessed the CDSS before completing a questionnaire. They used a retrospective chart review, specifically a half-day of pre-recorded patient card sheet data, to evaluate the CDSS. On the basis of pre-recorded cases, the goal was to evaluate how well the CDSS performed in identifying referrals and locally treatable cases that were actually made. The health care professionals then filled out questionnaires to provide their assessments and feedback on the CDSS. Each evaluation questionnaire was completed based on a half-day of ANC, pregnant patient, and PNC cases, as well as the health care professional's observation of the CDSS reaction to the presented cases. Next, the health care professionals answered a series of 5-point Likert scale items (1=strongly disagree, 2=disagree, 3=neutral, 4=agree, 5=strongly agree) about the CDSS [20]. The measurement instrument consisted of 22 parameters adopted from Ji et al's [12]

XSL•FO
RenderX

evaluation framework. The 22 measurement items were classified into 6 factors: system quality, information quality, service quality, perceived ease of use, user acceptability, and perceived benefits. Furthermore, we automated the questionnaire submission, which was accessed via a mobile phone or a laptop. Electronic questionnaire submission was preferred over paper-based alternatives. However, paper-based questionnaire submissions were used in some cases.

The CDSS at the Point of Care

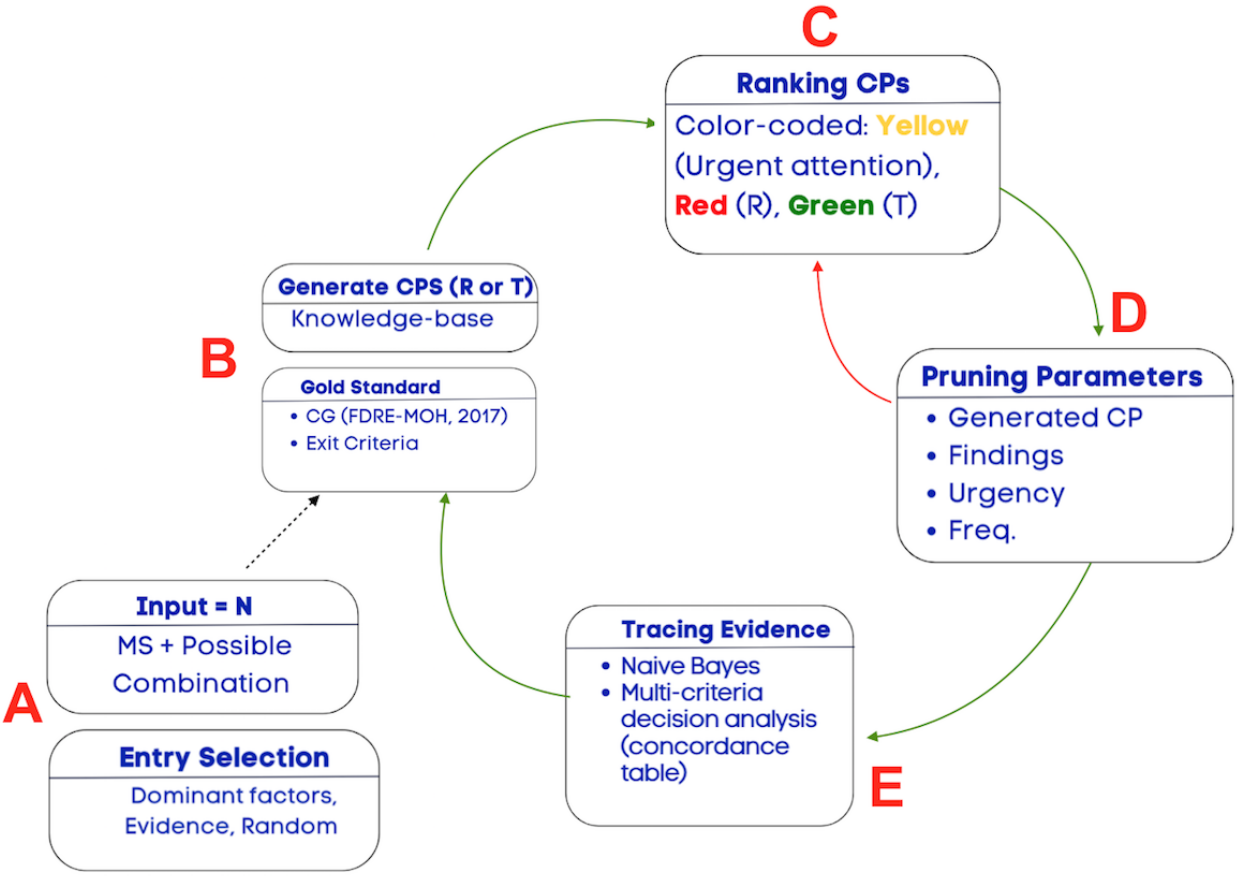
We designed and developed a CDSS to meet the requirements of LRSs. An intelligent clinical wizard, minimum data and data readiness, adaptable features, and low-cost infrastructure are some of the notable requirements and prerequisites of LRSs based on our previous results [7]. Our CDSS incorporates both existing knowledge-based guidelines and data-driven evidence to provide the most relevant information for frontline workers at the time of care delivery [11]. The CDSS provides CPs (or workflows) for point-of-care services. The CP is a critical component of a CDSS for identifying referral and locally treatable cases, which is delivered in the form of a concordance table for multicriteria decision analysis and output [11].

The CDSS has the following major goals:

- Delivering automated CPs and computer-assisted pruning and selection.
- Going beyond existing paper-based evidence that is noninteractive and challenging to grasp, the computerized CDSS was designed to be interactive for ease of use and optimal usage.
- Combining existing CGs and historical evidence (data-driven evidence) to generate an adaptable clinical workflow.

To get the most out of services, the CDSS provides an automated, interactively adaptable CP (or workflow). To reduce arbitrariness in entry point selection, the CDSS provides a range of choices for initiating the CP, such as using evidence from historical records, dominant factors, or randomly initiating the signs and symptoms based on CGs. Figure 2A presents additional information about entry point processing.

Figure 2. CP-processing workflow. CG: Clinical guideline; CP: Clinical pathway; FDRE-MOH: Federal Democratic Republic of Ethiopia Ministry of Health; Freq.: Frequency; MS: Measured symptom.



The process is interactive, and our algorithm uses measured symptoms (MSs) and a combination of MSs to process the CPs:

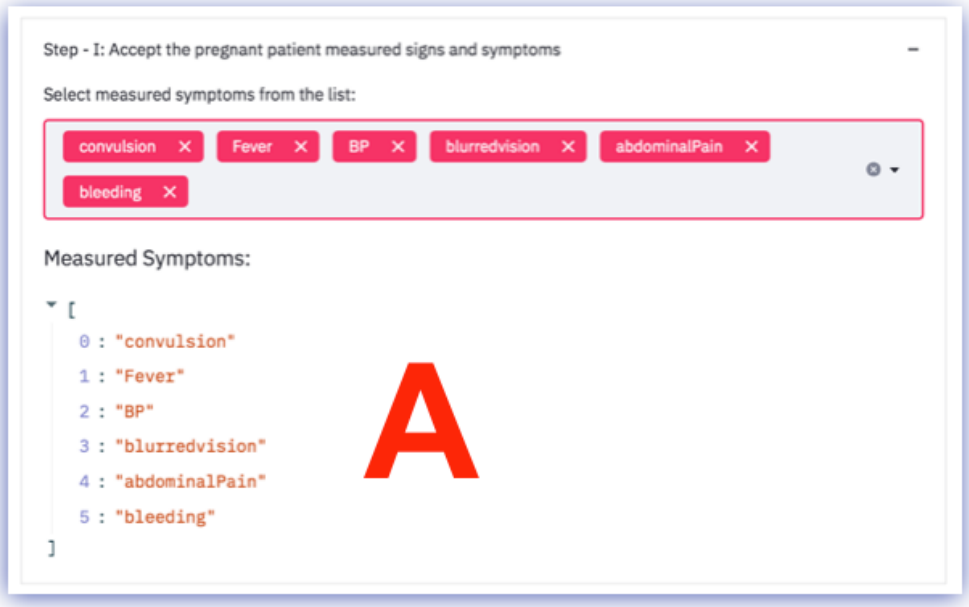
- First, all CPs based on the first MS are generated, as shown in Figure 2B. CGs are used as the gold standard and criterion for validating the generated CP (also referred to as an exit criterion). If the generated CP is already found on the generated list, the frequency counter is incremented. Otherwise, the generated CP is added (or appended) to the generated list of CPs. Federal Democratic Republic of Ethiopia Ministry of Health version 2017 (FDRE-MOH 2017) is used for CP processing.
- Second, a ranking of CPs is conducted to identify "referral" and "locally treatable" cases. The ranking is color-coded, as shown in Figure 2C, and the ranking criteria are based on CGs.

XSL•FO
RenderX

- Third, the dynamic CP list is pruned, as shown in Figure 2D. CP pruning is based on pruning parameters. If the generated CP list is empty, fall-back and adjustment of the pruning criterion are supported. The pruning process was designed to be interactive, flexible, responsive, and engaging. The user intervention allows for fine-tuning based on domain knowledge and provides trust and understanding for the health care professional. Pruning can also be based on findings if the health care professional requires pruning of specific CP findings. The findings are based on the CGs.
- Fourth, the naive Bayes algorithm and historical records are used to provide data-driven evidence, as shown in Figure 2E. The output is displayed in an easy-to-understand format, using a table to present the evidence. The ranked table provides evidence for assessing various factors, such as symptoms, findings, urgency, CP, CP frequency, accuracy, and prior and posterior probability, to facilitate evidence-based decision-making by the user. Since it provides evidence for analyzing various factors, we refer to it as multicriteria decision analysis. In further detail, the multicriteria output used for decision analysis is displayed in the form of a table, also known as a concordance table. A concordance table is a data (evidence) table used as a cross-reference for integrating evidence from many sources for decision support. In this study, it was used primarily for tracing what evidence was available to support the presented case and identifying the evidence's source (historical records or knowledge-based evidence). A more detailed step-by-step description of the algorithms is found in Figure 2A-E.
- Finally, the preceding steps are repeated for each additional MS.

In the end, the frontline worker must make the final decision based on the suggestions made by the algorithm. For this study and demonstration, the CDSS focused on 3 use cases, namely pregnant patient care, ANC, and PNC services. The sample user interface screenshot for each step is shown in Figures 3-7.

Figure 3. Screenshot of input processing. BP: blood pressure.



XSL•FO
RenderX

Figure 4. Screenshot of the generated CPs and the gold standard. ANC: antenatal care; CG: clinical guideline; CP: clinical pathway; KB: knowledge base; NC, not classified; PNC: postnatal care; R: referral; T: treatable.

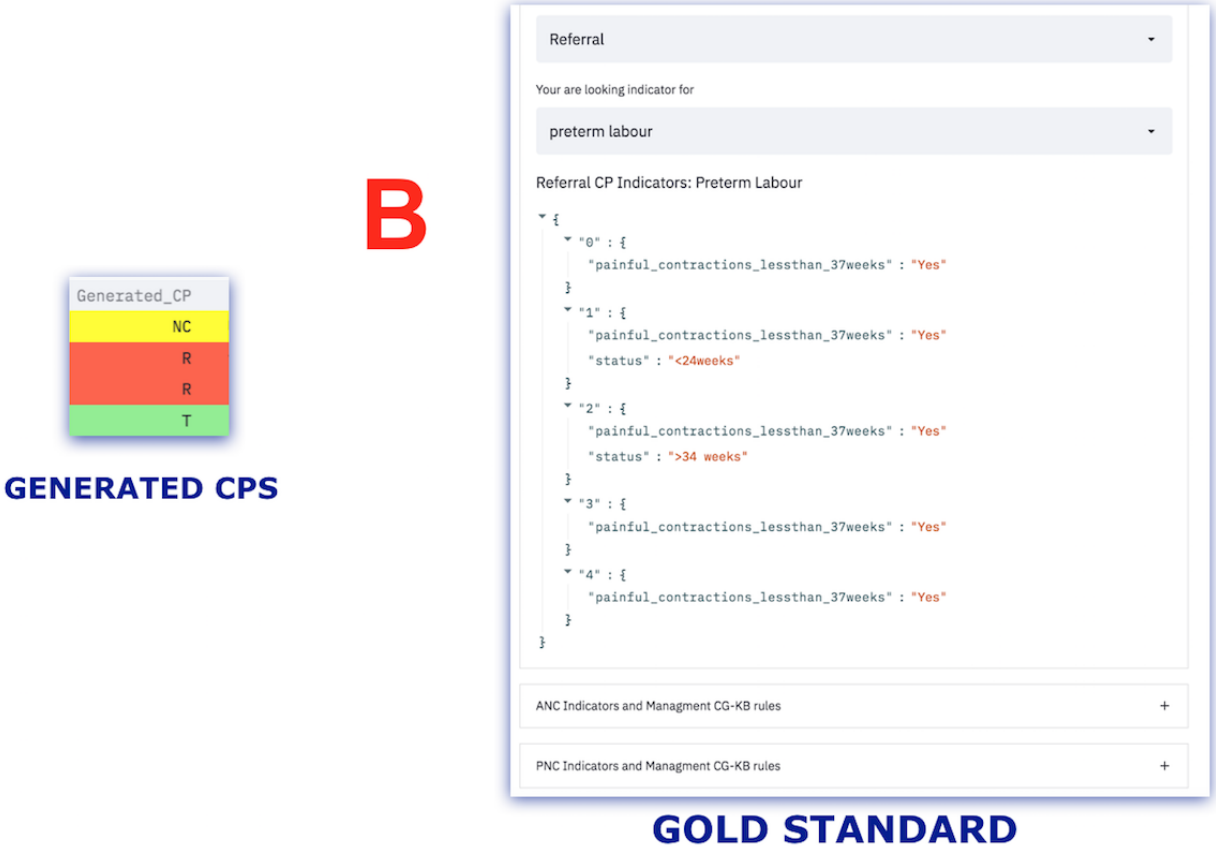
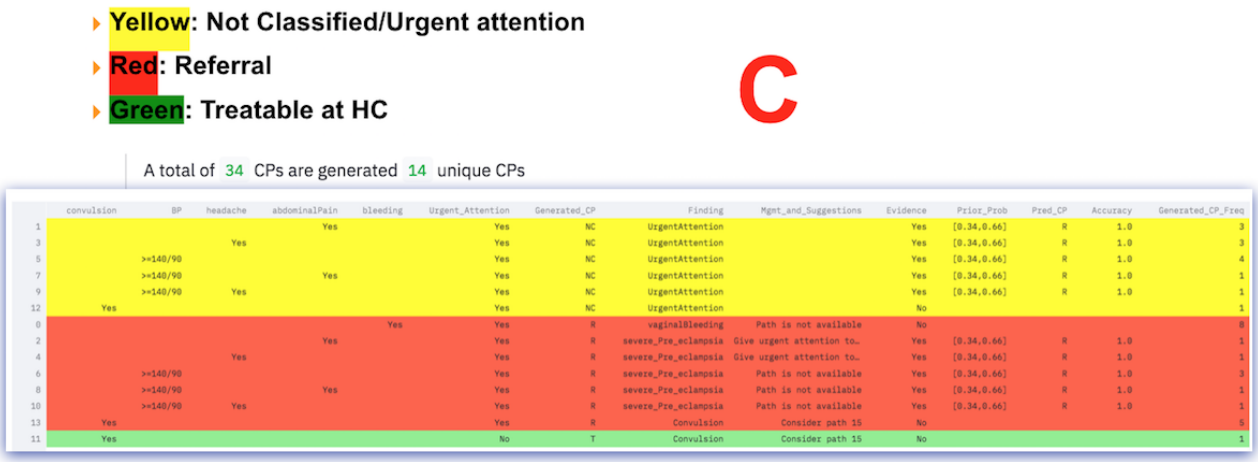


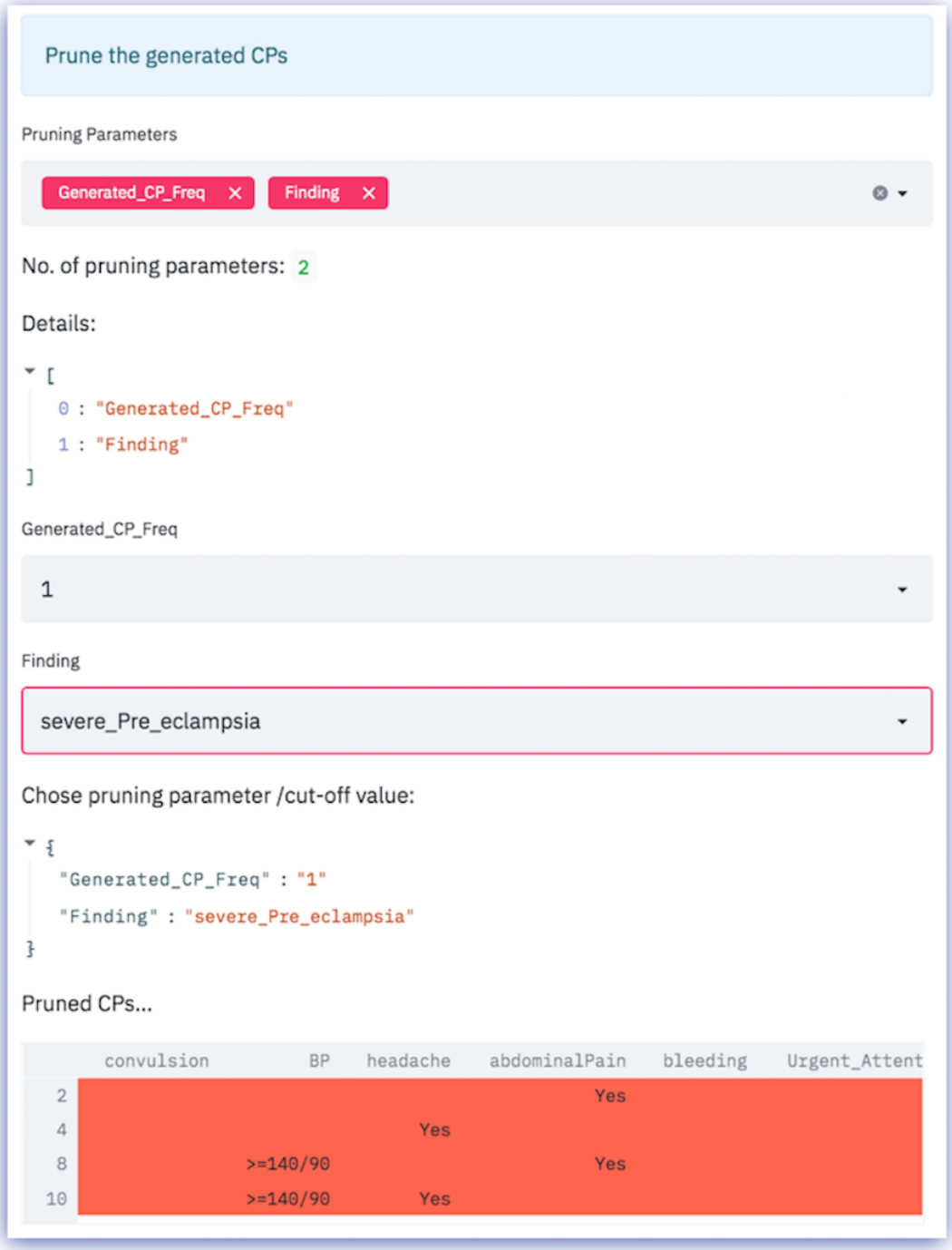
Figure 5. CP ranking. BP: blood pressure; CP: clinical pathway; HC: health center.



	convulsion	BP	headache	abdominalPain	bleeding	Urgent_Attention	Generated_CP	Finding	Mgmt_and_Suggestions	Evidence	Prior_Prob	Pred_CP	Accuracy	Generated_CP_Freq
1				Yes		Yes	NC	UrgentAttention		Yes	[0.34,0.66]	R	1.0	3
3			Yes			Yes	NC	UrgentAttention		Yes	[0.34,0.66]	R	1.0	3
5		>=140/90				Yes	NC	UrgentAttention		Yes	[0.34,0.66]	R	1.0	4
7		>=140/90		Yes		Yes	NC	UrgentAttention		Yes	[0.34,0.66]	R	1.0	1
9		>=140/90	Yes			Yes	NC	UrgentAttention		Yes	[0.34,0.66]	R	1.0	1
12	Yes					Yes	NC	UrgentAttention		No				1
0					Yes	Yes	R	vaginalBleeding	Path is not available	No				8
2				Yes		Yes	R	severe_Pre_eclampsia	Give urgent attention to…	Yes	[0.34,0.66]	R	1.0	1
4			Yes			Yes	R	severe_Pre_eclampsia	Give urgent attention to…	Yes	[0.34,0.66]	R	1.0	1
6		>=140/90				Yes	R	severe_Pre_eclampsia	Path is not available	Yes	[0.34,0.66]	R	1.0	3
8		>=140/90		Yes		Yes	R	severe_Pre_eclampsia	Path is not available	Yes	[0.34,0.66]	R	1.0	1
10		>=140/90	Yes			Yes	R	severe_Pre_eclampsia	Path is not available	Yes	[0.34,0.66]	R	1.0	1
13	Yes					Yes	R	Convulsion	Consider path 15	No				5
11	Yes					No	T	Convulsion	Consider path 15	No				1

XSL•FO
RenderX

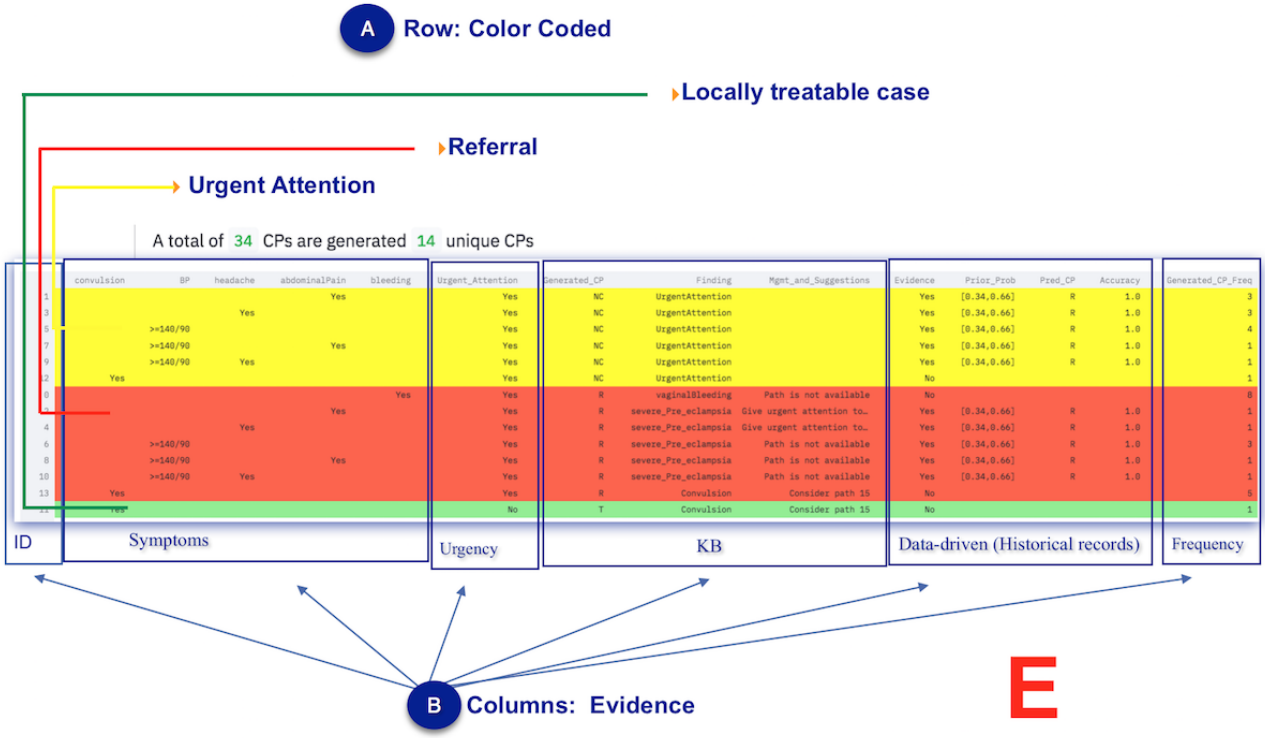
Figure 6. CP pruning. BP: blood pressure; CP: clinical pathway.



D

XSL•FO
RenderX

Figure 7. Concordance table. BP: blood pressure; CP: clinical pathway.



Data Analysis

Statistical Package for Social Sciences (SPSS; IBM Corporation) version 26.0 [21], Microsoft Excel [22], Python (version 3.7) [23], and SmartPLS (version 26.0) [24] were used to conduct the analysis and modeling.

We followed the procedures and recommendations of Boone and Boone [25] for the CDSS evaluation based on Likert data analysis. Latent variables were computed by summing the following items:

- The perceived ease of use was a latent variable based on learnability, operability, user interface, data entry, advice to display, and legibility items.
- Response time and stability items were used to assess system quality.
- Information quality was based on security and CP performance items.
- Acceptance included usage, confirmation of expectations, overall quality satisfaction, overall satisfaction, and the intention to use items.
- Perceived benefits were created using decision change (change in order behavior, change in CP) and process change (effectiveness, overall usefulness, adherence to standards, medical quality, and user knowledge and skills) items.

To assess the scale of the CDSS evaluation data set, the validity of the measurement model was checked. Convergent validity was assessed using factor loading and average variance extracted (AVE), with a factor loading threshold of more than 0.70 and an AVE threshold of >0.50 [26,27]. In this study, items with factor loadings of less than 0.70 were candidates for deletion. The internal consistency and reliability of the CDSS evaluation measurement model were assessed using Cronbach α [28] and composite reliability. A recommended value of >.70 for Cronbach α and composite reliability was accepted. We used the heterotrait-monotrait ratio of correlations (HTMT) [26,29] to check the discriminant validity of the measurement model and determined whether the value was less than 0.90 and acceptable. Moreover, perceived benefits are formative second-order construct based on decision changes and process changes. Collinearity was checked to ensure that it did not have a negative impact on the higher-order-construct measurement model, and critical levels of collinearity less than 0.50 were acceptable in this study, as recommended by Hair et al [26].

Following that, item-level and construct-level analyses were performed. On the one hand, an item-level analysis of the CDSS in LRSs between the 2 health centers was conducted. A nonparametric independent-samples statistical test, such as the Mann-Whitney *U* test [30], was used to see whether the 2 health centers were significantly different at the item level. We used the Mann-Whitney *U* test because we could not assume normality in either group and the independent data set observation assumptions were fulfilled, which are preconditions for the use of nonparametric data analysis [31]. Furthermore, there were no significant results from the Shapiro-Wilk test [32] on the normality of our evaluation data set. The significance level used for the inferential statistics was P=.05 and a 95% CI level.

On the other hand, we followed the recommendation of Boone and Boone [25] to use Pearson correlation for construct-level (latent variable) correlation analysis. As a result, Pearson correlation [25,33] was used to examine the factors influencing the acceptance of the CDSS in LRSs and the interrelationships between construct factors. In particular, the relationship between system quality and perceived ease of use, information quality

XSL•FO
RenderX

and perceived ease of use, user acceptance and perceived benefits, and user acceptance and information and system quality were explicitly explored.

Finally, structural equation modeling (SEM) is a multivariate statistical analysis technique that is used to analyze structural relationships. It is described in the literature as combined factor analysis and regression analysis for discovering relationships between measured variables and latent constructs [34]. There is a debate on how effective it is to discover causation beyond correlation. In papers dealing with applications of the technique, it is commonly used to express a causal hypothesis in a context where there is semantic information available that supports the validity of the hypothesis or at least does not contradict it [12,14,35]. Our study was a pilot study, not a full, cross-sectional analysis, and it intended to promote the use of partial least squares structural equation modeling (PLS-SEM) [26]. PLS-SEM was used to model the acceptance of the CDSS in LRSs, particularly to model the relationship between the CDSS evaluation measured items and construct variables, as well as between multiple construct variables. We noticed that penalized likelihood estimation algorithms based on regularized structural equation modeling (RegSEM) [36,37] and PLS-SEM [26] were the best candidates for our modeling. We preferred PLS-SEM for the following reasons:

- The SmartPLS [26,38] partial least squares (PLS) algorithm was used to analyze the model's path weight, and it performed well in Mburu and Oboko's [14] study.
- The variation-based structural equation models do not impose a sample size [39] or normality of distribution constraints [26,38].

Overall, to construct the PLS-SEM model for the CDSS in LRSs, first, composite factor analysis was used to examine the validity of the measurement model, including reliability and validity analysis. The relationships in path models with latent variables were then evaluated using PLS-SEM path analysis and coefficients. Finally, the statistical significance of PLS-SEM results, such as path coefficients, outer weights, Cronbach α, and coefficient of determination (R^2) values, was determined using bootstrapping [26]. The bootstrapping settings were percentile bootstrap, 2-tailed test type, and significance level=.05.

Results

Characteristics

The 7 CDSS evaluators were all female (ie, n=4, 57%, from the Jimma Health Center and n=3, 43%, from the Jimma Higher Two Health Center), who worked as health care professionals (eg, midwives and nurses) in the health centers' maternal and childcare health service units. In total, 73 CDSS evaluation responses were recorded based on 116 ANC, pregnant patient care, and PNC cases (n=4, 5%, during the first evaluation period and n=69, 95%, during the second evaluation period). The response was 73 since the evaluation response was based on a summary of half-day cases rather than a case-by-case response. The average time for evaluating the CDSS and completing the questionnaire was 52.35 minutes, with the smallest and longest durations being 31 and 98 minutes, respectively. The Jimma Health Center accounted for 65.5% (76/116) cases, while the Jimma Higher Two Health Center accounted for 34.5% (40/116) cases. Furthermore, we observed that each health center handled 4-6 (3%-5%) cases per day on average. Overall, the first round of evaluation lasted 2 days and included 18 ANC, pregnant patient care, and PNC cases in the Jimma Health Center, which is above average. In round 2, there were 75 ANC cases, 7 pregnant patient care cases, and 16 PNC cases during our evaluation period. The second round of evaluation took place in both health centers, the Jimma Health Center and the Jimma Higher Two Health Center.

The computer-aided CDSS evaluation's mean (SD) score ranged from 4.29 (SD 0.485) to 4.52 (SD 0.503). Table 1 provides more extensive details of each item score.

XSL•FO
RenderX

Table 1. Mean Likert scale scores and reliability analysis for computer-aided CDSS[a] evaluation in LRSs[b].

Construct and items	Value, minimum (maximum)	Score of 73 CDSS evaluation responses, mean (SD)
Perceived ease of use		
Learnability	2 (5)	4.30 (0.545)
Operability	3 (5)	4.29 (0.485)
User interface	3 (5)	4.34 (0.533)
Data entry	3 (5)	4.40 (0.571)
Advice display	3 (5)	4.37 (0.589)
Legibility	1 (5)	4.29 (0.905)
System quality		
Response time	3 (5)	4.38 (0.543)
Stability	2 (5)	4.38 (0.615)
Information quality		
Security	3 (5)	4.32 (0.550)
CP[c] performance	3 (5)	4.37 (0.540)
Decision change		
Change in order behavior	2 (5)	4.08 (0.640)
Change in CP	2 (5)	4.23 (.613)
Process changes		
Effectiveness	3 (5)	4.25 (0.494)
Overall usefulness	3 (5)	4.23 (0.635)
Adherence to standards	3 (5)	4.33 (0.502)
Medical quality	3 (5)	4.29 (0.612)
User knowledge and skills	2 (5)	4.30 (0.570)
Acceptance		
Usage	3 (5)	4.49 (0.580)
Confirmation of expectations	2 (5)	4.34 (0.628)
Satisfaction with overall quality	3 (5)	4.40 (0.571)
Overall satisfaction	__[d]	4.30 (0.570)
Intention to use	4 (5)	4.52 (0.503)

[a]CDSS: clinical decision support system.

[b]LRS: low-resource setting.

[c]CP: clinical pathway.

[d]Not applicable.

CDSS Evaluation Measurement Model

The factor loading of 20 (91%) of 22 items was greater than 0.70. The remaining items, legibility and medical quality, were eliminated since their factor loading value was less than 0.70. All the constructs had Cronbach α values greater than .70, except information quality, for which Cronbach α was .699, which is close to .70. Table 2 provides more information about the measurement model's construct reliability and validity.

To establish discriminant validity, the HTMT on construct factors was used, and the results showed that all constructs passed the test. Table 3 displays the results of the discriminant validity assessment.

XSL•FO
RenderX

Table 2. CDSS[a] measurement model's construct reliability and validity.

Construct and items	Convergent validity		Internal consistency and reliability	
	Factor loading (>0.70)	AVE[b] (>0.50)	Composite reliability (>0.70)	Cronbach α (>.70)
Perceived ease of use		0.588	0.847	.825
Learnability	0.721	—[c]	—	—
Operability	0.738	—	—	—
User interface	0.746	—	—	—
Data entry	0.856	—	—	—
Advise to display	0.836	—	—	—
System quality		0.879	0.869	.863
Response time	0.934	—	—	—
Stability	0.944	—	—	—
Information quality		0.763	0.767	.699
Security	0.825	—	—	—
CP[d] performance	0.930	—	—	—
Decision changes		0.776	0.712	.712
Change in order behavior	0.856	—	—	—
Change in CP	0.856	—	—	—
Process changes		0.650	0.824	.819
Effectiveness	0.773	—	—	—
Overall usefulness	0.813	—	—	—
Adherence to standards	0.896	—	—	—
User knowledge and skills	0.762	—	—	—
Acceptance		0.654	0.871	.867
Usage	0.738	—	—	—
Confirmation of expectations	0.819	—	—	—
Satisfaction with overall quality	0.806	—	—	—
Overall satisfaction	0.846	—	—	—
Intension to use	0.815	—	—	—
Perceived benefits				
Constructed based on decision and process changes	—	0.511	0.848	0.839

[a]CDSS: clinical decision support system.

[b]AVE: average variance extracted.

[c]Not applicable.

[d]CP: clinical pathway.

XSL•FO
RenderX

Table 3. CDSS[a] discriminant validity assessment.

Constructs	Perceived ease of use	Information quality	Perceived benefits	Perceived user acceptance	System quality
Perceived ease of use	__[b]	—	—	—	—
Information quality	0.855	—	—	—	—
Perceived benefits	0.643	0.852	—	—	—
Perceived user acceptance	0.616	0.877	0.877	—	—
System quality	0.545	0.839	0.618	0.779	—

[a]CDSS: clinical decision support system.

[b]Not applicable.

CDSS Evaluation Between the 2 Health Centers

The results of the nonparametric Mann-Whitney *U* test based on the 5-point Likert item evaluation data set collected from the Jimma Health Center and the Jimma Higher Two Health Center revealed that the 2 health centers did not differ significantly in the CDSS item-level evaluation factors, except stability (U=470.5, P=.022), overall usefulness (U=451.0, P=.012), adherence to standards (U=483, P=.024), confirmation of expectations (U=488.5, P=.04), satisfaction with overall quality (U=400.5, P=.001), and overall satisfaction (U=474.5, P=.023). The findings of the CDSS evaluation using the Mann-Whitney *U* test are shown in Table 4.

XSL•FO
RenderX

Table 4. Mann-Whitney U test results (*P*<.05).

Construct and items	Mean rank		Test statistics[a]	
	Jimma Health Center (n=42)	Jimma Higher Two Health Center (n=31)	Mann-Whitney *U*	Asymptotic significance (2-tailed) *P* value
Perceived ease of use				
Learnability	36.92	37.11	647.5	.962
Operability	38.75	34.63	577.5	.309
User interface	37.48	36.35	631.0	.794
Data entry	37.85	35.85	615.5	.653
Advice display	36.14	38.16	615.0	.650
Legibility	34.76	40.03	557.0	.249
System quality				
Response time	33.48	41.77	503.0	.057
Stability	32.70	42.82	470.5	.022
Information quality				
Security	35.49	39.05	587.5	.409
CP[b] performance	35.65	38.82	594.5	.466
Decision changes				
Change in order behavior	36.48	37.71	629.0	.767
Change in CP	35.51	39.02	588.5	.416
Process changes				
Effectiveness	35.82	38.60	601.5	.489
Overall usefulness	32.24	43.45	451.0	.012
Adherence to standards	33.00	42.42	483.0	.024
Medical quality	34.43	40.48	543.0	.174
User knowledge and skills	34.50	40.39	546.0	.164
Acceptance				
Usage	32.79	42.71	474.0	.024
Confirmation of expectations	33.13	42.24	488.5	.04
Satisfaction with overall quality	31.04	45.08	400.5	.001
Overall satisfaction	32.80	42.69	474.5	.023
Intension to use	35.38	39.19	583.0	.381

[a]Grouping variable: health center.

[b]CP: clinical pathway.

CDSS Evaluation Agreement Score Observation in the Jimma Health Center

Although the total number of observations in the first and second rounds of the CDSS evaluation were not equal, we found a positive mean agreement score increment in the majority of evaluation parameters at the Jimma Health Center, which was calculated using “agree” and “strongly agree” responses. Adherence to the standards agreement score, however, declined from 1.00 to 0.974. The first and second round CDSS evaluation agreement score observations at the Jimma Health Center are shown in Table 5.

XSL•FO
RenderX

Table 5. First and second round CDSS[a] evaluation agreement score observations at the Jimma Health Center.

Construct and items	Round 1: agreement score based on n=4 observations, mean (SD)		Round 2: agreement score based on n=38 observations, mean (SD)	
	Item level	Construct level	Item level	Construct level
Perceived ease of use	__[b]	0.708	—	0.982
Learnability	0.750	—	1.000	—
Operability	0.750	—	1.000	—
User interface	0.750	—	1.000	—
Data entry	0.750	—	1.000	—
Advice to display	0.750	—	1.000	—
Legibility	0.500	—	0.890	—
System quality	—	0.625	—	0.987
Response time	0.750	—	1.000	—
Stability	0.500	—	0.974	—
Information quality	—	0.750	—	0.974
Security	0.750	—	0.947	—
CP[c] performance	0.750	—	1.000	—
Decision changes	—	0.750	—	0.960
Change in order behavior	0.500	—	0.947	—
Change in CP	1.000	—	0.974	—
Process changes	—	0.800	—	0.953
Effectiveness	0.750	—	1.000	—
Overall usefulness	0.750	—	0.842	—
Adherence to standards	1.000	—	0.974	—
Medical quality	0.750	—	0.947	—
User knowledge and skills	0.750	—	0.974	—
User acceptance	—	0.750	—	0.958
Usage	0.750	—	0.947	—
Confirmation of expectations	0.500	—	0.974	—
Satisfaction with overall quality	0.750	—	0.947	—
Overall satisfaction	0.750	—	0.921	—
Intention to use	1.000	—	1.000	—

[a]CDSS: clinical decision support system.

[b]Not applicable.

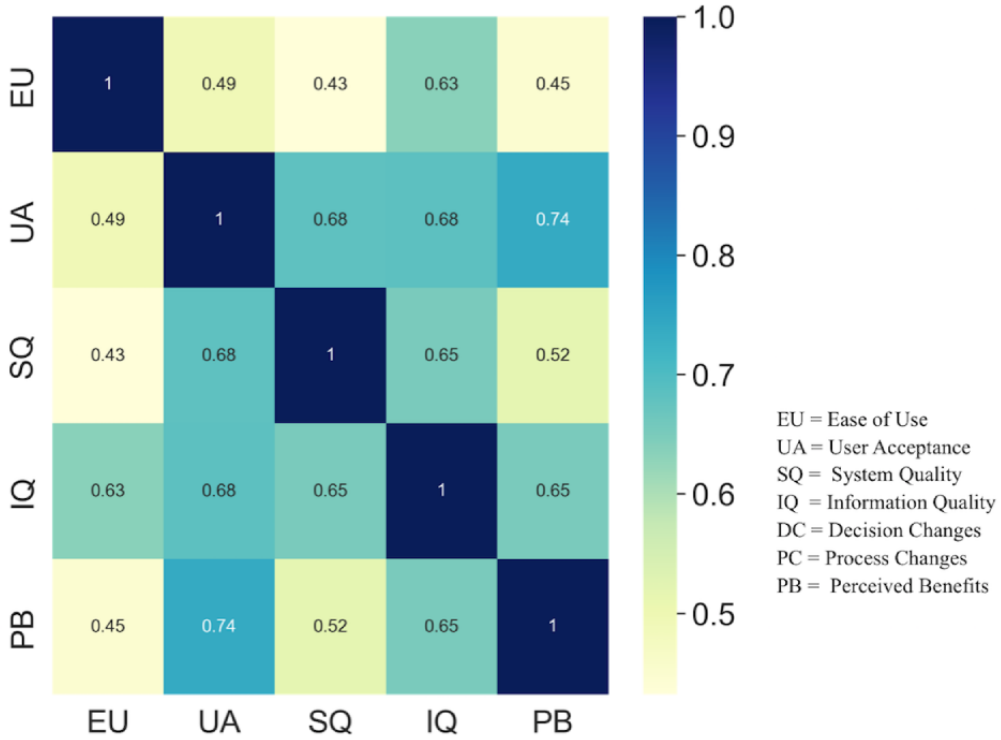
[c]CP: clinical pathway.

CDSS Evaluation: Construct Factor Interrelationships

We found a significant correlation (r =0.74) between user acceptance and perceived benefits, with perceived benefits as construct factors based on process changes and decision changes. The coefficient of correlation between perceived ease of use and information quality was r=0.63. User acceptance was also correlated with information quality and system quality, with r=0.68. Figure 8 depicts a more detailed Pearson correlation test result.

XSL•FO
RenderX

Figure 8. Pearson correlation (N=73). Correlation values ranging from 0.50 to 0.70 are considered moderate, from 0.70 to 0.90 are considered strong, and from 0.9 to 1.0 are considered very strong [33].



Modeling the Acceptance of the CDSS in LRSs

The perceived user acceptance coefficient of determination (R^2) was 0.703, showing that user acceptance is influenced by system quality, information quality, perceived ease of use, and perceived benefits (Figure 9). More precisely, system quality (β=.321, *P*<.05) and perceived benefits (β=.486, *P*<.05) were shown to have a significant influence. However, the perceived ease of use had no positive effect on CDSS acceptance (β=.015, *P*=.91). Information quality also had no positive effect on CDSS acceptance in this study (β=.149, and *P*=.25).

Furthermore, we found that the perceived ease of use was influenced by system quality, and information quality, with an R^2 value of 0.479. The path coefficient of information quality on the perceived ease of use was β=.021(*P*=.89), and hence, no significant effect was found. The path coefficient of system quality on the perceived ease of use was β=.678 (*P*<.05), that is, a significant influence, whereas the perceived benefits impacted by decision and process changes had an R^2 value of 0.983. The path coefficients of decision changes and process changes were β=.374 and β=.749, respectively, and were significant (*P*<.05). Figure 9 depicts the path weights, *P* values, and coefficient of determination (R^2) for the CDSS evaluation PLS-SEM model developed using the CDSS Jimma Health Center and Jimma Higher Two Health Center evaluation data sets. The results, shown in Figure 9, can be interpreted as perceived ease of use -> perceived user acceptance (β=.015 and *P*=.91), for example. Overall, we found that the perceived ease of use had no positive effect on CDSS acceptance (β=.015, *P*=.91) but, rather, on the system quality (β=.321, *P*<.05) and perceived benefits (β=.486, *P*<.05) of the CDSS. Further information is presented in Table 6.

XSL•FO
RenderX

Figure 9. Computer-aided CDSS evaluation PLS-SEM model generated from the computer-aided CDSS Jimma Health Center and the Jimma Higher Two Health Center evaluation data sets, showing path weights (β), *P* values, and coefficient of determination (R^2). The yellow boxes represent indicators (or parameters). The construct variables are represented by the circle. The path indicates the path weight and *P* value. For example, a 0.321 (.002) value from system quality -> perceived user acceptance shows that β=.321 and *P*=.002. CDSS: clinical decision support system; PLS-SEM: partial least squares structural equation modeling.

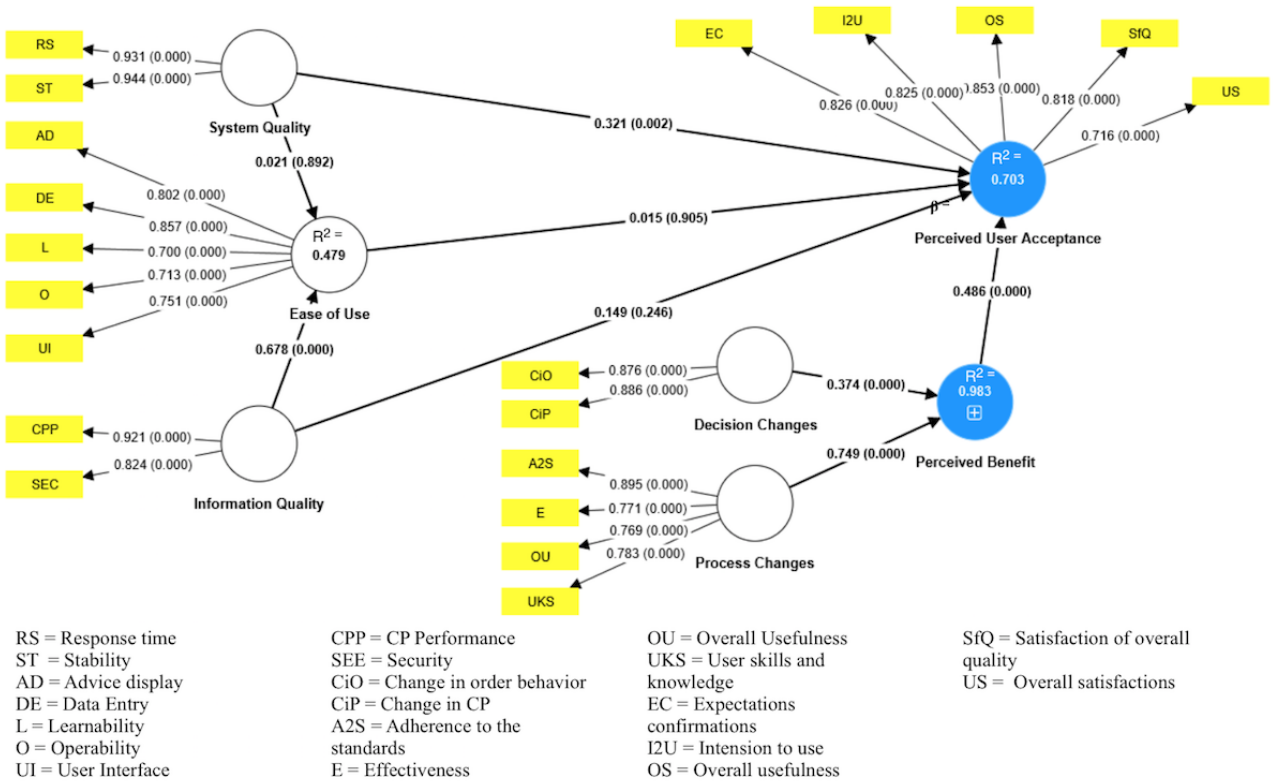


Table 6. Hypotheses conclusion based on PLS-SEM[a] findings (β=.015, *P*=.91).

Hypothesis	Path and relationships	PLS-SEM findings[b]			Conclusion
		β (SD)	*t* Statistics	*P* value	
Hypothesis (H)1: The perceived ease of use has a positive effect on the acceptance of a CDSS[c] in LRSs[d].	Perceived ease of use -> acceptance	.015 (0.123)	0.119	.91	Rejected
H2: System quality has a positive effect on the acceptance of a CDSS in LRSs.	System quality -> acceptance	.321(.102)	3.139	.002	Accepted
H3: Information quality has a positive effect on the acceptance of a CDSS in LRSs.	Information quality -> acceptance	.149 (0.128)	1.162	.25	Rejected
H4: Information quality has a positive effect on the perceived ease of use of a CDSS in LRSs.	Information quality -> perceived ease of use	.678 (0.122)	5.558	<.001	Accepted
H5: System quality has a positive effect on the perceived ease of use of a CDSS in LRSs.	System quality -> perceived ease of use	.021 (0.153)	.135	.89	Rejected
H6: Perceived benefits have a positive effect on the acceptance of a CDSS in LRSs.	Perceived benefits -> acceptance	.486 (0.115)	4.234	<.001	Accepted

[a]PLS-SEM: partial least squares structural equation modeling.

[b]Relationships were significant at *P*<.05.

[a]CDSS: clinical decision support system.

[d]LRS: low-resource setting.

XSL•FO
RenderX

Discussion

Principal Findings

This study aimed to evaluate a CDSS for use at the point of care in primary care LRSs. The health care professionals in this study evaluated user acceptance of the CDSS.

The Cronbach α scale of 22 items appeared to be internally consistent, exceeding the minimum value of .70 required for acceptable reliability [26-28,32]. In this study, the 2 health centers did not differ significantly in terms of the CDSS's perceived ease of use, information quality, and perceived benefits (decision changes and process changes). However, we found a significant difference in system quality, such as stability, and perceived user acceptance, such as overall usefulness, adherence to standards, confirmation of expectations, satisfaction with overall quality, and overall satisfaction. This variation could be attributed to the first round of evaluation, which was based on the Jimma Health Center, or to the fact that more cases were observed in the Jimma Health Center than in the Jimma Higher Two Health Center, but more research and analysis are required. Furthermore, based on the first and second rounds of the CDSS evaluation, we observed a positive agreement score increment at the Jimma Health Center. However, this study was unable to observe a change in the Jimma High Two Health Center, since the first round of evaluation was limited to the Jimma Health Center.

This study highlighted a correlation between construct variables using Pearson correlation. The CDSS's system quality, information quality, and perceived benefits were vital for its acceptance in the LRS. The perceived benefits were based on decision and process changes. In accordance with our results, previous studies have demonstrated that the acceptance and use of mHealth apps in LRSs are influenced by users' perceptions that are aligned with health needs and expectations [14]. However, in this study, the perceived ease of use was moderately correlated with CDSS acceptance, whereas Ji et al [12] suggested that the perceived ease of use has a significant and direct impact on the acceptance of a CDSS. Figure 8 depicts further information about the Pearson correlation between the construct variables of the CDSS. Overall, we observed a low positive Pearson correlation between the perceived ease of use and acceptance, as well as between system quality and the perceived ease of use, when we considered the strength of correlation classification as in Mukaka [33]. System quality and acceptance, information quality and acceptance, and information quality and perceived ease of use all showed a moderately positive correlation. There was a high positive correlation between perceived benefits and acceptance, supporting Pande et al's [40] finding that perceived usefulness is significantly correlated to the intention to use.

This study also used PLS-SEM to evaluate several factors that impact the acceptance of a CDSS in LRSs. The result demonstrated that user acceptance is impacted by system quality, information quality, and perceived benefits, with an R^2 value of 0.703, as shown in Figure 9. The perceived benefits influenced by decision and process changes had an R^2 value of 0.983, whereas the R^2 score for the perceived ease of use as impacted by system and information quality was only 0.479. All retained R^2 values were greater than 0.10, as suggested by Falk and Miller [41]. The R^2 values of the perceived user acceptance and perceived benefits were substantial, as also indicated by the CDSS results of Cohen [18], who reported R^2>0.26, and Chin [42], who reported R^2>0.67. However, according to the criteria of Hair et al [43], R^2 of perceived benefits is greater than 0.75 and substantial, while R^2 of the perceived ease of use and user acceptance is greater than 0.50 and moderate. However, Mohamed et al [44] showed that the coefficient of determination must be larger than 0.19, the path coefficient between latent variables must be at least 0.1, and the significance level must be at least .05 in order to validate the model. Our CDSS evaluation model meets all these criteria, except the path coefficient from perceived ease of use to perceived user acceptance, which was 0.015. Hair et al [26], however, stated that path coefficients with standardized values greater than 0.20 are typically significant, while in this study, the path coefficient from perceived ease of use to user acceptance was 0.015, which is less than 0.10 and not significant. More information is depicted in Figure 9, which includes the details of the CDSS assessment PLS-SEM model developed from the CDSS Jimma Health Center and Jimma Higher Two Health Center evaluation data sets, including path weights, *P* values, and the coefficient of determination (R^2).

Overall, as shown in Table 6, the PLS-SEM results suggested that the perceived ease of use has no positive effect on CDSS acceptance (β=.015, *P*=.91) but, rather, on system quality (β=.321, *P*=.002) and perceived benefits (β=.486, *P*<.001) of the CDSS. We also observed that information quality had a positive influence on the perceived ease of use (β=.678, *P*<.001). However, system quality had no favorable impact on the perceived ease of use (β=.021, *P*=.89). The detailed conclusions and summary based on PLS-SEM are shown in Table 6.

Limitations

In this study, we evaluated our own proof-of-principle CDSS in LRSs. The small sample size and low number of cases in our study might limit the generalizability of our findings. As a result, difficulties that were not identified during this investigation may be identified during a longitudinal and case-by-case evaluation.

Conclusion

We designed and developed a CDSS based on LRS requirements, which we evaluated in 2 LRSs in Ethiopia: the Jimma Health Center and the Jimma Higher Two Health Center. Our overall result indicates that user acceptance is impacted by system quality, information quality, perceived ease of use, and perceived benefits, with an R^2 value of 0.703. Specifically, system quality and perceived benefits have a direct impact on user acceptance of the CDSS in LRSs. In this study, however, we found that the perceived ease of use and information quality had no positive effect on CDSS acceptability. Overall, the proposed acceptance model includes specific factors and

XSL•FO
RenderX

variables, which is an important step toward the successful adoption and implementation of a CDSS in LRSs.

Acknowledgments

We would like to thank the health care professionals at the Jimma Health Center's and the Jimma Higher Two Health Center's maternal and childcare health service units. Sisters Aynalem Wubet, Digiti Abaraya, and Asiya, in particular, deserve special appreciation. Their voluntarism will not be forgotten. The evaluation was possible due to their efforts in administering the clinical decision support system evaluation from beginning to end. We would also like to thank Mr Balew Ayalew (a Jimma University lecturer) for facilitating and assisting with the supervision of the evaluation at the Jimma Higher Two Health Center. Finally, we thank the Network for Advancement of Sustainable Capacity in Education and Research in Ethiopia (NASCERE) program.

Authors' Contributions

GST, DAS, GKT, FV, JC, and BJ conceptualized the research goals and objectives, as well as the methodology. GST conducted data curation, formal analysis, investigation, visualization, and drafting of the manuscript. DAS, GKT, FV, JC, and BJ were involved in the supervision, validation, visualization, and review and editing of the manuscript, as well as the final proofreading. All authors have read and approved the final manuscript.

Conflicts of Interest

None declared.

References

1. Haux R. Health information systems - past, present, future. Int J Med Inform 2006 Mar;75(3-4):268-281. [doi: 10.1016/j.ijmedinf.2005.08.002] [Medline: 16169771]
2. Azubuike M, Ehiri J. Health information systems in developing countries: benefits, problems, and prospects. J R Soc Promot Health 1999 Sep 07;119(3):180-184. [doi: 10.1177/146642409911900309] [Medline: 10518358]
3. Pellé KG, Rambaud-Althaus C, D'Acremont V, Moran G, Sampath R, Katz Z, et al. Electronic clinical decision support algorithms incorporating point-of-care diagnostic tests in low-resource settings: a target product profile. BMJ Glob Health 2020 Feb 28;5(2):e002067 [FREE Full text] [doi: 10.1136/bmjgh-2019-002067] [Medline: 32181003]
4. Abbas JJ, Smith B, Poluta M, Velazquez-Berumen A. Improving health-care delivery in low-resource settings with nanotechnology: challenges in multiple dimensions. Nanobiomedicine (Rij) 2017 Mar 29;4:1849543517701158 [FREE Full text] [doi: 10.1177/1849543517701158] [Medline: 29942391]
5. López DM, Rico-Olarte C, Blobel B, Hullin C. Challenges and solutions for transforming health ecosystems in low- and middle-income countries through artificial intelligence. Front Med (Lausanne) 2022 Dec 2;9:958097 [FREE Full text] [doi: 10.3389/fmed.2022.958097] [Medline: 36530888]
6. Alliance for Health Policy and Systems Research. Primary Health Care Systems (PRIMASYS): Case Study from Ethiopia; abridged version. Geneva: World Health Organization; 2017:1-12.
7. Tegenaw G, Amenu D, Ketema G, Verbeke F, Cornelis J, Jansen B. Using clinical guidelines and card sheets for guiding the design of data-driven clinical pathways. J Health Inform Dev Ctries 2021 Nov 12;15(2):1-18.
8. Tegenaw GS, Amenu D, Ketema G, Verbeke F, Cornelis J, Jansen B. Analysis of low resource setting referral pathways to improve coordination and evidence-based services for maternal and child health in Ethiopia. PLoS One 2022 Aug 25;17(8):e0273436 [FREE Full text] [doi: 10.1371/journal.pone.0273436] [Medline: 36007079]
9. Biru A, Birhan D, Melkamu G, Gebeyehu A, Omer AM. Pathways to improve health information systems in Ethiopia: current maturity status and implications. Health Res Policy Syst 2022 Jun 29;20(1):78 [FREE Full text] [doi: 10.1186/s12961-022-00860-z] [Medline: 35768819]
10. Tegenaw G, Amenu D, Ketema G, Verbeke F, Cornelis J, Jansen B. Design approaches for executable clinical pathways at the point of care in limited resource set-tings to support the clinical decision process: review of the state of the art. 2021 Presented at: MobiHealth 2021: International Conference on Wireless Mobile Communication and Healthcare; November 13-14, 2021; Chongqing, PRC. [doi: 10.1007/978-3-031-06368-8_13]
11. Tegenaw G, Amenu D, Ketema G, Verbeke F, Cornelis J, Jansen B. A hybrid approach for designing dynamic and data-driven clinical pathways point of care instruments in low resource settings. 2021 Presented at: MEDINFO 2021: One World, One Health—Global Partnership for Digital Innovation: 18th World Congress on Medical and Health Informatics; October 2-4, 2021; Virtual p. 316. [doi: 10.3233/shti220087]
12. Ji M, Genchev GZ, Huang H, Xu T, Lu H, Yu G. Evaluation framework for successful artificial intelligence–enabled clinical decision support systems: mixed methods study. J Med Internet Res 2021 Jun 02;23(6):e25929 [FREE Full text] [doi: 10.2196/25929] [Medline: 34076581]

XSL•FO
RenderX

13. Tegenaw GS, Amenu D, Ketema G, Verbeke F, Cornelis J, Jansen B. Evaluating a clinical decision support point of care instrument in low resource setting. BMC Med Inform Decis Mak 2023 Mar 30;23(1):51 [FREE Full text] [doi: 10.1186/s12911-023-02144-0] [Medline: 36998074]
14. Mburu S, Oboko R. A model for predicting utilization of mHealth interventions in low-resource settings: case of maternal and newborn care in Kenya. BMC Med Inform Decis Mak 2018 Jul 17;18(1):67. [doi: 10.1186/s12911-018-0649-z] [Medline: 30016943]
15. Chin W, Newsted P. Structural equation modeling analysis with small samples using partial least squares. In: Hoyle RE, editor. Statistical Strategies for Small Sample Research. Thousand Oaks, CA: SAGE Publications; 1999:307-341.
16. Cohen J. A power primer. Psychol Bull 1992 Jul 01;112(1):155-159. [doi: 10.1037//0033-2909.112.1.155] [Medline: 19565683]
17. Goodhue D, Lewis W, Thompson R. PLS, small sample size, and statistical power in MIS research. 2006 Presented at: HICSS'06: 39th Annual Hawaii International Conference on System Sciences; January 4-7, 2006; Kauai, Hawaii p. 202b. [doi: 10.1109/hicss.2006.381]
18. Cohen J. Statistical Power Analysis for the Behavioral Sciences. Abingdon-on-Thames, UK: Routledge; 1988.
19. Sideridis G, Simos P, Papanicolaou A, Fletcher J. Using structural equation modeling to assess functional connectivity in the brain: power and sample size considerations. Educ Psychol Meas 2014 Mar 04;74(5):733-758. [doi: 10.1177/0013164414525397]
20. Bertram D. Likert scales: CPSC 681—topic report. Poincare. 2006. URL: http://poincare.matf.bg.ac.rs/~kristina/topic-dane-likert.pdf [accessed 2024-05-09]
21. SPSS Statistics 26.0.0. IBM Corp. URL: https://www.ibm.com/docs/en/spss-statistics/26.0.0 [accessed 2024-05-09]
22. Microsoft Excel. Microsoft Corporation. 2022. URL: https://office.microsoft.com/excel [accessed 2024-05-09]
23. Van Rossum G, Drake F. Python 3 Reference Manual. Scotts Valley, CA: CreateSpace; 2009.
24. Sarstedt M, Cheah J. Partial least squares structural equation modeling using SmartPLS: a software review. J Market Anal 2019 Jun 27;7(3):196-202. [doi: 10.1057/s41270-019-00058-3]
25. Boone H, Boone D. Analyzing Likert data. JOE 2012 Apr 1;50(2):48. [doi: 10.34068/joe.50.02.48]
26. Hair Jr JF, Hult GT, Ringle CM, Sarstedt M. A Primer on Partial Least Squares Structural Equation Modeling (PLS-SEM). In: A Primer on Partial Least Squares Structural Equation Modeling (PLS-SEM). Thousand Oaks, CA: SAGE Publications; 2013:211-213.
27. Vinzi V, Chin W, Henseler J, Wang H. Handbook of Partial Least Squares. Berlin: Springer; 2010.
28. Gliem JA, Gliem RR. Calculating, interpreting, and reporting Cronbach's alpha reliability coefficient for Likert-type scales. 2003 Presented at: Midwest Research-to-Practice Conference in Adult, Continuing, and Community Education; 2003; Columbus, OH.
29. Henseler J, Ringle CM, Sarstedt M. A new criterion for assessing discriminant validity in variance-based structural equation modeling. J Acad Mark Sci 2014 Aug 22;43(1):115-135. [doi: 10.1007/s11747-014-0403-8]
30. DePuy V, Berger V, Zhou Y. Wilcoxon-Mann-Whitney test: overview. In: Wiley StatsRef: Statistics Reference Online. Hoboken, NJ: John Wiley & Sons; 2014:14.
31. Grech V, Calleja N. WASP (write a scientific paper): parametric vs. non-parametric tests. Early Hum Dev 2018 Aug;123:48-49. [doi: 10.1016/j.earlhumdev.2018.04.014] [Medline: 29678516]
32. Hendrickson L, Cronbach LJ. Designing evaluations of educational and social programs. In: Contemporary Sociology. San Francisco, CA: Jossey-Bass; 1983:309.
33. Mukaka MM. Statistics corner: a guide to appropriate use of correlation coefficient in medical research. Malawi Med J 2012 Sep;24(3):69-71 [FREE Full text] [Medline: 23638278]
34. Bollen K, Pearl J. Eight myths about causality and structural equation models. In: Handbook of Causal Analysis for Social Research. Dordrecht: Springer Netherlands; 2013:301-328.
35. Jung SY, Hwang H, Lee K, Lee H, Kim E, Kim M, et al. Barriers and facilitators to implementation of medication decision support systems in electronic medical records: mixed methods approach based on structural equation modeling and qualitative analysis. JMIR Med Inform 2020 Jul 22;8(7):e18758 [FREE Full text] [doi: 10.2196/18758] [Medline: 32706717]
36. Huang P, Chen H, Weng L. A penalized likelihood method for structural equation modeling. Psychometrika 2017 Jun 17;82(2):329-354. [doi: 10.1007/s11336-017-9566-9] [Medline: 28417228]
37. Jacobucci R, Grimm KJ, McArdle JJ. Regularized structural equation modeling. Struct Equ Model 2016 Apr 12;23(4):555-566 [FREE Full text] [doi: 10.1080/10705511.2016.1154793] [Medline: 27398019]
38. Hair JF, Risher JJ, Sarstedt M, Ringle CM. When to use and how to report the results of PLS-SEM. Eur Bus Rev 2019 Jan 14;31(1):2-24. [doi: 10.1108/ebr-11-2018-0203]
39. Rosseel Y. Small sample solutions for structural equation modeling. In: Small Sample Size Solutions: A Guide for Applied Researchers and Practitioners. Abingdon-on-Thames, UK: Routledge; 2020:226-238.
40. Pande T, Saravu K, Temesgen Z, Seyoum A, Rai S, Rao R, et al. Evaluating clinicians' user experience and acceptability of LearnTB, a smartphone application for tuberculosis in India. Mhealth 2017 Jul 27;3:30 [FREE Full text] [doi: 10.21037/mhealth.2017.07.01] [Medline: 28828377]
41. Falk R, Miller N. A Primer for Soft Modeling. Akron, OH: University of Akron Press; 1992.

XSL•FO
RenderX

42. Chin WW. The partial least squares approach to structural equation modeling. In: Marcoulides GA, editor. Modern Methods for Business Research. Mahwah, NJ: Lawrence Erlbaum Associates; 1998:295-336.
43. Hair JF, Ringle CM, Sarstedt M. PLS-SEM: indeed a silver bullet. J Mark Theory Pract 2014 Dec 08;19(2):139-152. [doi: 10.2753/mtp1069-6679190202]
44. Mohamed Z, Ubaidullah N, Yusof S. An evaluation of structural model for independent learning through connectivism theory and Web 2.0 towards students' achievement. 2018 Presented at: ICASE 2018: International Conference on Applied Science and Engineering; October 6-7, 2018; Sukoharjo, Indonesia p. 1-5. [doi: 10.2991/icase-18.2018.1]

Abbreviations

ANC: antenatal care
AVE: average variance extracted
CDSS: clinical decision support system
CG: clinical guideline
CP: clinical pathway
FDRE-MOH: Federal Democratic Republic of Ethiopia Ministry of Health
HTMT: heterotrait-monotrait ratio of correlations
LRS: low-resource setting
mHealth: mobile health
MS: measured symptom
PLS: partial least squares
PLS-SEM: partial least squares structural equation modeling
PNC: postnatal care
SEM: structural equation modeling

Edited by A Kushniruk; submitted 28.03.23; peer-reviewed by M Thomson, N Diouf; comments to author 23.07.23; revised version received 11.09.23; accepted 20.01.24; published 11.06.24.

Please cite as:
Tegenaw GS, Sori DA, Teklemariam GK, Verbeke F, Cornelis J, Jansen B
Evaluation of a Computer-Aided Clinical Decision Support System for Point-of-Care Use in Low-Resource Primary Care Settings: Acceptability Evaluation Study
JMIR Hum Factors 2024;11:e47631
URL: https://humanfactors.jmir.org/2024/1/e47631
doi:10.2196/47631
PMID:38861298

©Geletaw Sahle Tegenaw, Demisew Amenu Sori, Girum Ketema Teklemariam, Frank Verbeke, Jan Cornelis, Bart Jansen. Originally published in JMIR Human Factors (https://humanfactors.jmir.org), 11.06.2024. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on https://humanfactors.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Assessing the Usability and Feasibility of Digital Assistant Tools for Direct Support Professionals: Participatory Design and Pilot-Testing

Patrice D Tremoulet[1], PhD; Andrea F Lobo[2], PhD; Christina A Simmons[1], PhD; Ganesh Baliga[2], PhD; Matthew Brady[2]

[1]Department of Psychology, Rowan University, Glassboro, NJ, United States

[2]Department of Computer Science, Rowan University, Glassboro, NJ, United States

Corresponding Author:
Patrice D Tremoulet, PhD
Department of Psychology
Rowan University
201 Mullica Hill Rd
Robinson Hall Room 115K
Glassboro, NJ, 08028
United States
Phone: 1 8562564500 ext 53777
Email: tremoulet@rowan.edu

Abstract

Background: The United States is experiencing a direct support professional (DSP) crisis, with demand far exceeding supply. Although generating documentation is a critical responsibility, it is one of the most wearisome aspects of DSPs' jobs. Technology that enables DSPs to log informal time-stamped notes throughout their shift could help reduce the burden of end-of-shift documentation and increase job satisfaction, which in turn could improve the quality of life of the individuals with intellectual and developmental disabilities (IDDs) whom DSPs support. However, DSPs, with varied ages, levels of education, and comfort using technology, are not likely to adopt tools that detract from caregiving responsibilities or increase workload; therefore, technological tools for them must be relatively simple, extremely intuitive, and provide highly valued capabilities.

Objective: This paper describes the development and pilot-testing of a digital assistant tool (DAT) that enables DSPs to create informal notes throughout their shifts and use these notes to facilitate end-of-shift documentation. The purpose of the pilot study was to assess the usability and feasibility of the DAT.

Methods: The research team applied an established user-centered participatory design process to design, develop, and test the DAT prototypes between May 2020 and April 2023. Pilot-testing entailed having 14 DSPs who support adults with IDDs use the first full implementation of the DAT prototypes during 2 or 3 successive work shifts and fill out demographic and usability questionnaires.

Results: Participants used the DAT prototypes to create notes and help generate end-of-shift reports. The System Usability Scale score of 81.79 indicates that they found the prototypes easy to use. Survey responses imply that using the DAT made it easier for participants to produce required documentation and suggest that they would adopt the DAT if this tool were available for daily use.

Conclusions: Simple technologies such as the DAT prototypes, which enable DSPs to use mobile devices to log time-stamped notes throughout their shift with minimal effort and use the notes to help write reports, have the potential to both reduce the burden associated with producing documentation and enhance the quality (level of detail and accuracy) of this documentation. This could help to increase job satisfaction and reduce turnover in DSPs, both of which would help improve the quality of life of the individuals with IDDs whom they support. The pilot test results indicate that DSPs found the DAT easy to use. Next steps include (1) producing more robust versions of the DAT with additional capabilities, such as storing data locally on mobile devices when Wi-Fi is not available; and (2) eliciting input from agency directors, families, and others who use data about adults with IDDs to help care for them to ensure that data produced by DSPs are relevant and useful.

(JMIR Hum Factors 2024;11:e51612) doi:10.2196/51612

XSL•FO
RenderX

KEYWORDS

technology prototype; data collection; documentation; direct support professionals; intellectual and developmental disabilities; pilot test; mobile phone

Introduction

Background

In 2019, more than 2 million adults with intellectual and developmental disabilities (IDDs) were living in the United States [1]. Many rely upon direct support professionals (DSPs) for assistance with activities of daily living, such as hygiene, dressing, taking medications properly, eating, accessing and navigating stores, learning vocational skills, participating in therapeutic activities, and socializing [2]. It is widely understood that the quality of life of adults with IDDs is significantly impacted by the quality of the support they receive from DSPs [3-11]. Unfortunately, there has been a shortage in the DSP workforce for more than a decade [7,12]. This shortage and a DSP turnover rate of 44.8% in 2017 led the President's Committee for People with Intellectual Disabilities to declare a crisis in the direct support workforce [13]. The DSP shortage and high turnover rate, each of which is associated with reduced quality of life for those served by DSPs [3,7], have both significantly increased since the COVID-19 pandemic [14]. Therefore, it is not surprising that the rate of COVID-19 infection among adults with IDDs was disproportionally high and that their quality of life decreased between 2019 and 2020 [15].

DSPs have indicated that generating required documentation is one of the least rewarding and most onerous aspects of their jobs [16]. Several factors make this task challenging, including difficulty recalling details of work performed many hours ago; fatigue; being rushed because employers require them to check out on a time clock at specific times to avoid overtime pay; fear of being accused of copying and pasting content from previous shifts; frequent interruptions by clients or other staff members; and, for some, challenges in writing in a nonnative language [16]. High turnover increases the communication demands on DSPs, including the need to generate detailed documentation to help bring new DSPs up to speed on their clients' needs. However, overworked DSPs who prioritize clients' medical and behavioral health needs may struggle to find time to document and share all relevant data during shift changes, unintentionally leaving their clients vulnerable to medical errors and inadequate support [14].

Technology enabling in-the-moment data collection can increase the efficiency of direct support staff and raise the quality of the data they produce, both of which can help improve care for their clients [17-19]. However, most data collection and reporting tools used to capture data about individuals with IDDs are targeted toward providers who work one-on-one with children [16] (eg, paraprofessionals or behavior technicians). These tools are not appropriate for DSPs who support multiple adults with IDDs during their shifts. In addition, the responsibilities of DSPs are much wider in scope than those of paraprofessionals and behavior technicians, who are typically not responsible for tasks such as administering medication and helping prepare meals. DSPs' responsibilities are closer to those of home health workers who provide older adults or other adults with medical assistance and help with activities of daily living. Many high-income countries are already facing shortages of in-home health workers, while demand for them, as well as for DSPs, is expected to grow in many countries [20-22].

Recognizing that DSPs and the adults with IDDs who depend upon them could both potentially benefit from improved data collection and documentation, the research team used an established user-centered design methodology called Interaction Design and Engineering for Advanced Systems (IDEAS) [23] to design and pilot-test a suite of technology components called a digital assistant tool (DATs) to support data collection and reporting. This methodology relies upon frequent input and feedback from the target users, in this case, DSPs, to ensure that novel technology solutions will be useful, usable, and accepted by the DSPs [23].

Objectives

The primary purpose of the pilot study described here was to assess the usability of the DAT prototypes. The main objectives of the pilot study were as follows:

1. Identify design inconsistencies and usability problem areas of the DAT.
2. Observe representative users interacting with the DAT prototypes to help assess whether this technology could be effective, efficient, and well received by DSPs.
3. Establish baseline performance and user satisfaction levels in anticipation of more widespread testing of improved versions of the DAT.

Methods

Study Design

Overview

This project applied the IDEAS methodology, a user-centered participatory design approach that relies upon frequent input and feedback from target users to ensure that novel technology solutions will be useful, usable, and accepted by the users [23]. There are 6 steps in the IDEAS process: needs analysis, requirements generation, design and engineering, interface review, implementation, and evaluation. The first phase of this project included the first 2 steps; the second phase included the third, fourth, and fifth steps; and the third phase, which is the focus of this paper, comprises the last step.

Phase 1: Needs Analysis and Requirements Generation

To conceptualize the potential use of technology by DSPs, the research team sought to understand their perspectives on current data collection and documentation techniques and their ideas on how digital technology could be applied to support their work. The results of this exploratory descriptive research, which included focus groups, ethnographic observations, and a survey,

XSL•FO
RenderX

are described elsewhere [19]. Using the findings of this formative research, the team developed a list of design principles for our first set of prototypes, shown in Textbox 1.

Textbox 1. Summary of formative research results and the corresponding design principles for the initial digital assistant tool (DAT) prototypes.

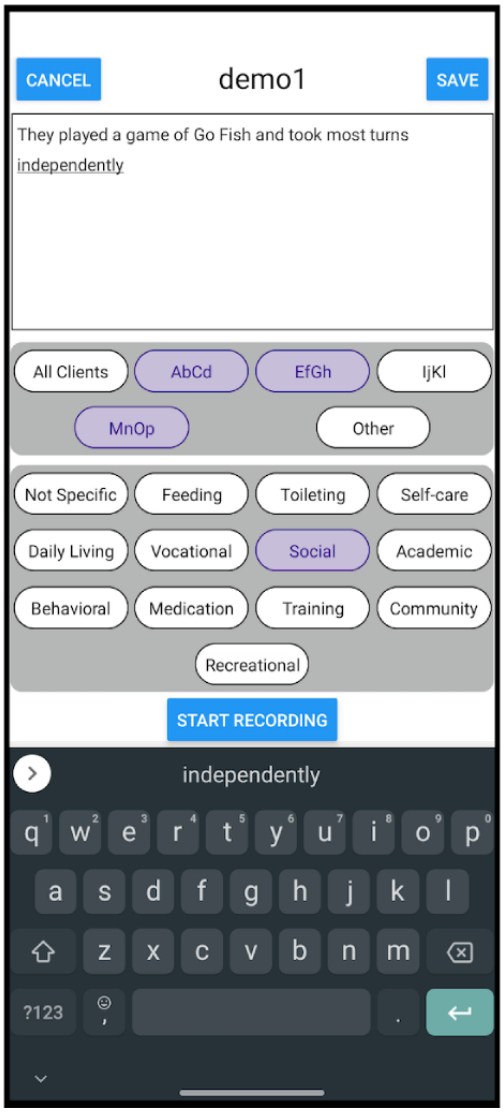
- Direct support professionals (DSPs) need to track and remember large amounts of information about multiple clients.
  - The DAT should automatically store notes, it should allow users to associate notes with specific clients, and it should provide access to notes on demand.
- DSPs must continuously monitor clients for safety while tracking behaviors.
  - The DAT should enable users to quickly and easily create notes while attending to clients, it should run on mobile and wearable devices, and it should not require a lengthy authentication process.
- DSPs must not allow clients to recognize when DSPs are creating notes about them.
  - The DAT should enable DSPs to create notes unobtrusively.
- DSPs are comfortable using smartphones but desire simple, easy-to-use data logging capabilities.
  - The DAT should prioritize usability—its user interfaces should be very simple, DAT log-ins should not time out during a shift, and the DAT should provide readable language.
- DSPs do not want employers to be able to access their notes.
  - The DAT should feature high levels of security and privacy and allow only DSPs to see their notes.
- DSPs must provide either chronological or categorically organized reports.
  - The DAT should time-stamp notes, it should allow users to associate notes with a topic or category, and it should allow users to sort and filter notes.
- DSPs must not copy text from prior days' reports into the current report.
  - The DAT should allow users to copy time-stamped notes into a clipboard or Word document to use to help write reports.

Phase 2: Design and Engineering, Interface Review, and Implementation

Once the research team had decided on the initial set of capabilities for the DAT, they designed a suite of technology prototypes that (1) enable quick and easy in-the-moment data collection; and (2) allow DSPs to access a private, secure web portal to review, sort, filter, and organize their notes to facilitate end-of-shift documentation. This suite includes 4 components: a mobile app that currently runs on Android smartphones; a private, secure web application that allows DSPs to access, review, and organize notes that they created with the mobile application; a cloud-based center that houses the data; and an administrative website for creating and managing user accounts. As these components were being architected, researchers shared user interface design concepts for the mobile app and the web portal with DSPs to obtain their feedback. The initial wireframes for the mobile app included multiple screens that would allow users not only to create notes but also to review and edit them (Multimedia Appendix 1). After 4 iterations, the team settled upon a very simple single-screen note creation design for the mobile app, also called the Note Creation App; and a spreadsheet-like view of saved notes for the web application, also referred to as the Note Review App (both of which were implemented; refer to Figures 1-3).

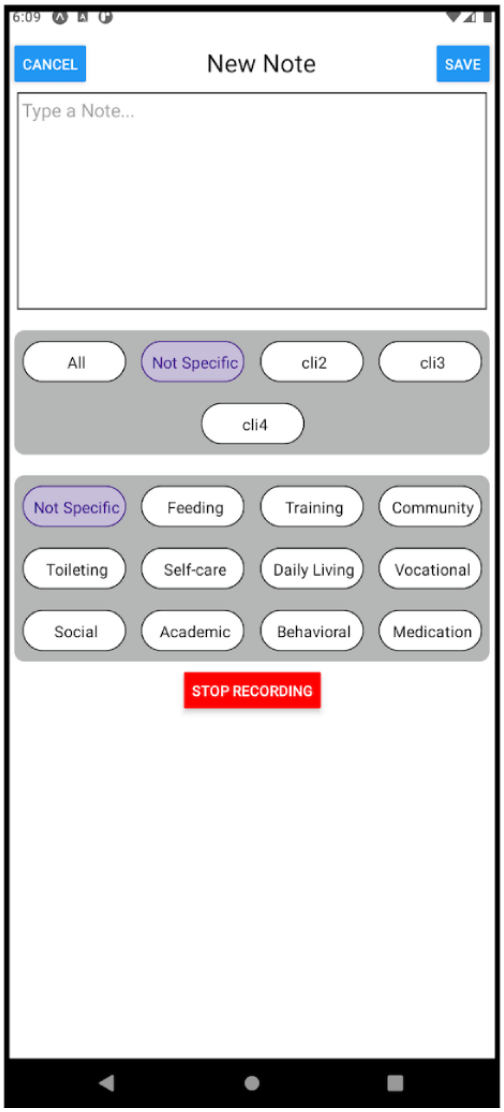
XSL•FO
RenderX

Figure 1. The mobile app, used to create informal notes. This screen shows the touchscreen keyboard used to type new notes.



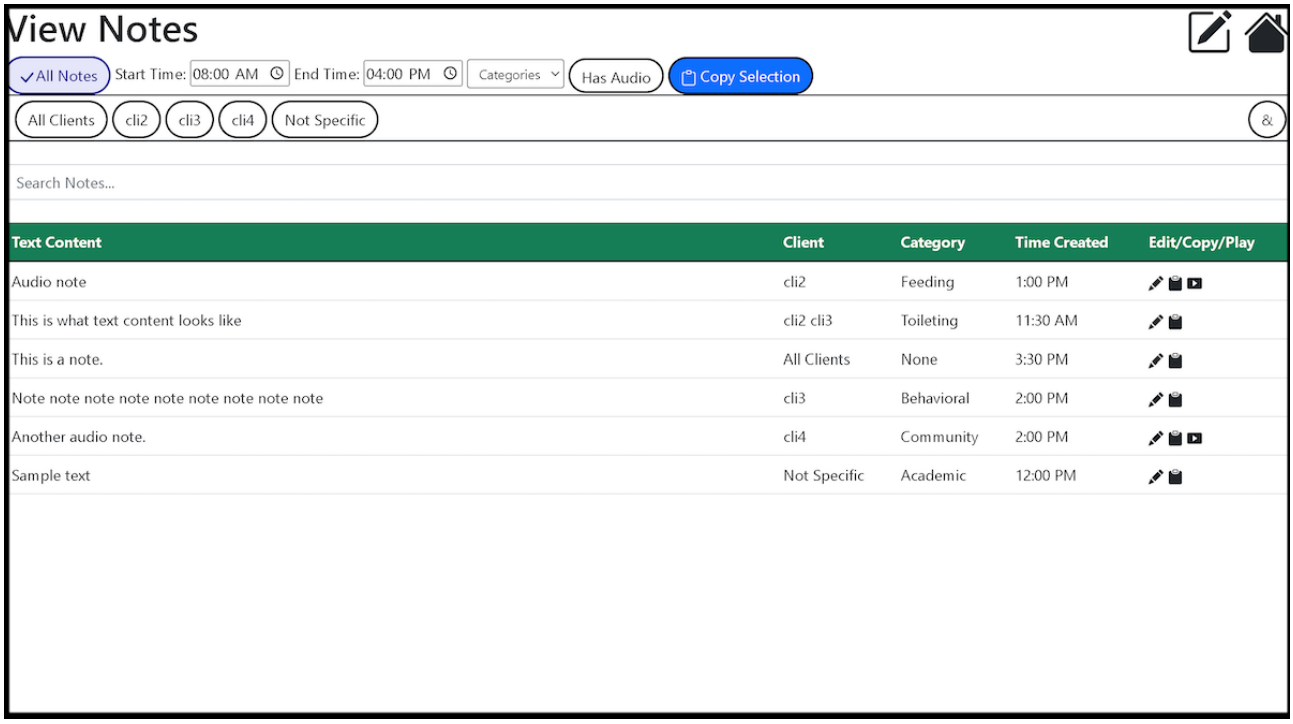
XSL•FO
RenderX

Figure 2. The mobile app, used to create informal notes. This screen shows the screen used for voice-based note creation.



XSL•FO
RenderX

Figure 3. The web application, used to organize notes and copy data into report generation software. cli2: client 2; cli3: client 3; cli4: client 4.



Phase 3: Evaluation

Once the initial prototypes were implemented, engineering and psychology undergraduate students tested them in laboratory settings; the engineering students focused on performance testing, and the psychology students focused on providing usability feedback. After all known bugs had been fixed, and usability improvements had been implemented, the team began scheduling pilot tests.

Participants and Recruitment

Overview

The pilot study was intended to determine *proof of concept*; therefore, the number of participants was not determined based on traditional power analysis calculations. Guidance on how many participants to include in pilot studies varies, with some recommending between 10 and 30 [24,25] and others suggesting 12 [26,27]. Meanwhile, the System Usability Scale (SUS), the instrument used to measure baseline usability, requires at least 8 to 10 participants to produce reliable results [28]. The research team worked with 3 DSP service provider partners to recruit 16 participants with the goal of having at least 12 (75%) of them test the prototypes during multiple shifts.

The service providers' senior management members notified their staff via email and during staff meetings that members of the research team would be bringing technology tools designed to be used by DSPs on specific dates. All staff who were working to support adults with IDDs in the designated program for 3 consecutive shifts during the research team visits were eligible to serve as pilot test participants. Due to variability in testing sites, including staff characteristics, the level of support needed by clients served by the staff, and the timing of the staff's work shifts, demographic data and results are reported separately for each test site.

Site 1 (Day Program)

The first pilot-testing site was a day program for adults with IDDS where DSPs and clients were each assigned to 1 of 5 different rooms. Of the 10 DSPs who worked in this setting on the first day of testing, 6 (60%) were selected to serve as pilot testers; 2 (20%) were excluded because they were not scheduled to work 3 shifts in a row in the same setting, and 2 (20%) were willing but unable to participate because there were only 6 smartphones with the DAT mobile app installed available; thus, the first 6 DSPs who volunteered were enrolled. All 6 participants were native English speakers; 5 (83%) were women. Of the 6 participants, 2 (33%) were aged between 18 and 24 years, 1 (17%) was aged between 25 and 34 years, 1 (17%) was aged between 35 and 44 years, and 2 (33%) were aged between 45 and 54 years. All identified as Black or African American. Of the 6 participants, 1 (17%) had a high school degree, while the remaining 5 (83%) had taken some college classes but did not have college degrees. Experience working with adults with IDDs ranged from 1.5 to >20 years, with an average of 8 years, 9.6 months. The amount of time they had held their positions ranged from 4 months to 3 years. Due to last-minute schedule changes, of the 6 participants, 2 (33%) used the DAT prototypes for 2 consecutive shifts, whereas the other 4 (67%) used them for 3 consecutive shifts.

Site 2 (Private Home)

The second test site was a private home that housed a single adult with a high level of support needs. Of the 12 different staff members who worked in this home, 4 (33%) were eligible to participate, based on their schedules, and all volunteered to serve as pilot testers. All 4 participants from this site were women and native English speakers; 1 (25%) was aged between 18 and 24 years, and the remaining 3 (75%) were aged between 25 and 34 years. Of the 4 participants, 1 (25%) identified as Black or African American, 1 (25%) as Hispanic, and 2 (50%)

XSL•FO
RenderX

as White. Of the 4 participants, 1 (25%) had a high school degree, 2 (50%) had some college but no degree, and 1 (25%) had a college degree. Their experience working with individuals with IDDs ranged from 7 months to 7 years, with an average of 3 years, 11.4 months. The amount of time they had been working in the private home ranged from 3 months to 1.5 years.

Site 3 (Provider-Managed Group Homes)

The third test site included 3 group homes, all located in the same county, that housed 3 to 4 adults with IDDs. Each of the 2 staff members in each of the 3 homes volunteered to serve as pilot testers. Of the 6 volunteers, 2 (33%) had work schedule changes that prevented them from using the DAT prototypes after their first shift and were removed from the study; the remaining 4 (67%) volunteers were included in the study. All 4 study participants were women and native English speakers. Of the 4 participants, 3 (75%) identified as Black Hispanic, and 1 as *other*; 1 (25%) had a college degree, and 3 (75%) had high school degrees and no college experience. Of the 4 participants, 1 (25%) was aged between 18 and 24 years, 1 (25%) was aged between 25 and 34 years, and 2 (50%) were aged between 45 and 54 years. The amount of experience they had working with adults with IDDs ranged from 1 to 22 years, with an average of 11 years, 9 months. The amount of time they had held their positions ranged from 1 to 22 years, with an average of 7 years, 4 months. Due to last-minute schedule changes, of the 4 participants, 2 (50%) used the DAT prototypes for 2 consecutive shifts, whereas 2 (50%) used them for 3 consecutive shifts.

Materials

Each participant was provided with an Android smartphone (Motorola Moto G Power 2021 running Android version 11 RZBS31.Q2-14327-25) that had the DAT mobile app preinstalled. For testing at sites 2 and 3, each smartphone also had shortcuts to all study surveys (demographic survey, post–shift 1 survey, and post–final shift survey) located on its main screen. Shortcuts were not provided on the smartphones during testing at site 1 because the research team was present to provide links in person at that site. All participants used employer-provided laptops to access the DAT web application at the end of their shifts.

Procedure

On the first day of testing, the principal investigator (PI) obtained informed consent from each participant. She explained that they would be awarded gift cards at the end of the multiday test trial that would be credited with US $75 per shift for their first 2 shifts and US $100 for their third shift. Next, she asked participants to use a computer to complete a short demographic questionnaire, administered through Qualtrics (Qualtrics International Inc; Multimedia Appendix 2).

While participants were filling out the demographic survey, the research team used the administrative website to create user accounts and location-based shifts, which included the initials of the clients whom each participant would be supporting and relevant data categories for the DSPs who worked in this setting (eg, feeding, toileting, behavior, medication, and social skills). These categories were identified by asking staff supervisors to select from a list of possible categories; they were also invited to add items not on the list.

After the participant finished the demographic survey, the researcher showed them the web application, which was populated with sample data previously entered by the research team. The researcher explained to the participant that this application would be used at the end of their shift to support writing the required reports. Subsequently, the researcher provided each participant with an Android smartphone, demonstrated how to use the mobile app, and invited the participants to try creating an audio note as well as a text-based note.

Participants were also invited to review, sort, and filter their sample notes using the web application during the training session. Once participants indicated that they knew how to use both the mobile app and the web application, they were given an opportunity to ask questions and informed that they could request help or address questions to the PI at any point during their shifts. They were then directed to use the smartphone that the research team had provided to log audio and text-based notes about their work while otherwise performing their job as usual throughout their shift.

Site 1 Testing

During pilot-testing in the day program, 2 to 3 members of the research team stayed on site during all 3 pilot-testing shifts. One research team member monitored the use of the mobile app through the administrative website, and 2 other members intermittently visited the rooms where the DSPs who were testing the DAT were assigned to work to observe or answer questions. In 1 room, the ratio of DSPs to clients was 1:3 on day 1 and 1:4 on days 2 and 3. In the second room, the ratio of DSPs to clients was 2:5 on day 1 and 2:4 on days 2 and 3. In the third room, the ratio was 2:3 on days 1 and 2 and 3:3 on day 3 (the third DSP in this room was in training and did not have access to the DAT). Finally, the fourth room, which was only included on day 3, had a ratio of 3:8, and only 1 DSP in this room used the DAT.

Throughout all their shifts, the participants used the mobile app to create short notes. They had the option to use voice or a touchscreen keyboard to enter text, and they could also use touchscreen buttons to associate notes with ≥1 clients or to indicate a specific category (eg, feeding, self-care, and medication).

Toward the end of the first day of pilot-testing, when participants were ready to start working on their required reports, a researcher helped them log in to the web application so that they could review and organize the notes that they had created during the shift. If needed, the researcher reviewed functionality, such as sorting, filtering, and copying 1 or multiple notes to a clipboard. The team then observed how the DSPs used the web application to help them create end-of-shift reports. Most of the DSPs (5/6, 83%) decided to paste their notes into either the clipboard or a text file and use 1 of these to draft the text portion of their end-of-shift report, rather than copying note content directly into the application that they use to submit their reports. Once the participants had finished their reports, they were

XSL•FO
RenderX

directed to use either the laptop used to create reports or the smartphone to fill out a web-based post–shift 1 survey (Multimedia Appendix 3). Participants were asked to rate 7 items, using a Likert scale ranging from 1 to 5, to indicate their level of agreement. Of the 7 items, 3 (43%) described their experience using the mobile app to log notes, their experience using the web application to review and organize the notes, and their opinions about the quality of the behavioral data and end-of-shift reports that they had provided that day. The items in this survey, which were also included in the final survey, were developed by the authors for this study and were not tested for validity or reliability.

At the end of days 2 and 3, the research team members were available to assist the participants, but all of them were able to access the web application and use it to help them with their reports independently. At the end of day 3, the participants were directed to fill out the final survey (Multimedia Appendix 4). The survey began with 17 items that were rated on a Likert scale ranging from 1 to 5 showing level of agreement. The first 7 items, which were about data collected and reports created, were the same as the first 7 statements rated after the first shift. The last 10 items were those included on the SUS survey, a well-established survey for assessing ease of use and user satisfaction [28,29]. Eleven open-ended questions followed the 17 statements. They asked participants to share their overall impressions of the mobile app and the web application, what they thought about the layout of the mobile app and the web application, what concerns they had about using the DAT, what it was like writing end-of-shift reports while using the DAT, and what additional feedback they wanted to share. These questions were designed to obtain feedback on existing functionality and elicit suggestions for changes or additions.

Once the final survey was submitted, the research team collected the smartphones, and the PI provided gift cards.

Site 2 Testing

As the private home housed a single adult, testing at this site was sequential, with only 1 (25%) of the 4 participants using the DAT prototypes on any given shift. In addition, only the PI visited this site when the DSPs who worked there participated in the pilot test. The PI met with each participant in the home approximately 1 hour before their shift started and obtained informed consent, provided the participant with a smartphone running the DAT mobile app, and taught them how to use the DAT mobile app and the DAT web application. Once training was completed, the PI placed shortcuts to the post–shift 1 survey as well as the final survey on the main screen of the smartphone and emailed the participant a private, secure link to the web application site where they could review and organize their notes. Finally, the PI answered any questions that the participant had, made sure the participant had a power cord and knew how to charge the smartphone with the mobile app, provided a phone number and email address to use in case questions or technical difficulties arose during their shifts, and departed. Although the research team members were not on site, they were able to monitor the use of the DAT mobile app and the DAT web application using the DAT administrative portal. All participants used the smartphone to create notes and accessed the web portal at the end of their shifts without support from the research team. Toward the end of each participant's first shift, the PI sent a reminder email that contained a link to the post–shift 1 survey (which was also accessible via a link on the smartphone). Toward the end of the participants' third shift, the PI sent an email to confirm a meeting time to deliver gift cards and pick up the smartphones; the email also contained a link to the final survey.

Site 3 Testing

Two members of the research team visited each of the 3 group homes during the first day of testing at site 3 to obtain informed consent, hand out the smartphones with the DAT mobile app, and conduct training. The research team members had placed shortcuts to the demographic, post–shift 1, and final surveys on the home screens of the smartphones in advance. After training was complete, the research team members helped participants to create a bookmark to conveniently access the web application on the computers that the participants used to produce their end-of-shift documentation. The PI also emailed links to the surveys and a secure link to access the web application so that the participants could easily access the surveys and their private website for reviewing and organizing their notes from their computers. Next, the researchers gave participants an email address and a phone number that could be used to reach the PI and encouraged them to use these if they had questions or concerns while using the DAT prototypes. The researchers then departed. During the rest of this shift, the research team members monitored the use of the DAT mobile app and the DAT web application via the administrative website. Toward the end of the first shift, the PI emailed participants reminders to fill out the post–shift 1 survey; and toward the end of the third shift, the PI emailed reminders to fill out the final survey. After the third shift had ended for all participants, the PI visited each of the group homes to pick up the smartphones and chargers and distribute gift cards.

Data Preparation and Analysis

Survey data were extracted from Qualtrics. For each site, descriptive statistics for demographic data were summarized; and the frequencies of the Likert scale responses, except for the 10 SUS items, were computed. The responses to the SUS items from all 14 participants were used to compute a single SUS score, using established computations [29]. The research team members also analyzed open-ended responses from all participants thematically, using inductive open coding [30]. Following the 6-step method formulated by Braun and Clarke [30] as explained by Maguire and Delahunt [31], 2 members of the research team categorized different response types for each question and used these to establish codes; next, these researchers and the PI collaboratively identified high-level themes that ran across multiple questions. Finally, the research assistants independently coded responses to all open-ended questions using the final code set.

In addition, in keeping with the exploratory nature of this work, the research team members analyzed the notes that were collected using the mobile app to understand how DSPs used this technology. In particular, the team reviewed the number of text-based and audio notes that each DSP produced during each

XSL•FO
RenderX

of their shifts, the number of words in each type of note, and the number of notes that were started but canceled (not saved) by each DSP during each shift.

Ethical Considerations

This research complies with the American Psychological Association code of ethics and was approved by the Rowan University Institutional Review Board (PRO-2020-001085). Informed consent was obtained from each participant on the first day of testing. Regarding compensation, at the end of the multiday test trial, participants were awarded gift cards that were credited with US $75 per shift for their first 2 shifts and US $100 for their third shift.

Results

Note Logs

Data logs from 34 DSP shifts were imported into Excel (Microsoft Corp) for analysis. Across these shifts, a total of 373 notes were saved. Another 41 notes were started, but they were canceled (not saved). Table 1 shows, for each site, the medians and IQRs of all saved notes per shift, audio notes saved per shift, text-based notes saved per shift, canceled notes per shift, notes created during a DSP's first pilot-testing shift, notes created during a DSP's second pilot-testing shift, notes created during a DSP's third pilot-testing shift, words per audio note, words per text note, and words per note.

Table 1. Types of direct support professional notes produced at each site during the pilot-testing.

Computation	Site 1, median (IQR)	Site 2, median (IQR)	Site 3, median (IQR)
Audio notes per shift	2 (1-9)	0 (0)	3 (2-5)
Text-based notes per shift	8 (4-14)	10 (7.5-14)	7 (4-12)
Total notes per shift	10 (5-12)	10 (7.5-14)	10 (2-16)
Canceled notes per shift	0 (0-1.5)	0 (0-3.5)	0 (0-1.5)
Notes created during first shift of pilot-testing	6 (4-12)	7.5 (6-9)	8.5 (5-12)
Notes created during second shift of pilot-testing	10 (5-12)	10 (5.5-14.5)	6 (2-10)
Notes created during third shift of pilot-testing	6 (8.5-12.5)	12.5 (6-16)	9.5 (7-12)
Words per audio note	9 (6-17)	N/A[a]	8.5 (8-14)
Words per text note	12 (6-22.5)	6.5 (3-9)	11 (8-15)
Words per note	10 (6-22)	6.5 (3-9)	10.5 (8-15)

[a]N/A: not applicable.

Post–Shift 1 Survey Ratings Questions

The post–shift 1 survey began with 7 statements that were rated on a scale ranging from 1 to 5 to show level of agreement (1=strongly disagree, 2=disagree, 3=neutral, 4=agree, and 5=strongly agree). Table 2 shows these statements as well as the frequencies of each of the ratings given to each statement.

XSL•FO
RenderX

Table 2. Frequencies of the post–shift 1 survey Likert scale responses for each test site. Response scale ranged from 1=strongly disagree to 5=strongly agree.

Statement	Site 1 (n=6), response; n (%)	Site 2 (n=4), response; n (%)	Site 3 (n=4), response; n (%)
I am confident that today's data sheets are accurate	• 4; 1 (17) • 5; 5 (83)	• 5; 4 (100)	• 3; 2 (50) • 4; 1 (25) • 5; 1 (25)
I found it easy to record behavior data for all clients today	• 4; 1 (17) • 5; 5 (83)	• 5; 4 (100)	• 3; 2 (50) • 4; 1 (25) • 5; 1 (25)
I believe today's behavior data will be valuable to others	• 4; 1 (17) • 5; 5 (83)	• 5; 4 (100)	• 3; 3 (75) • 4; 1 (25)
I found it easy to write session notes today	• 5; 6 (100)	• 3; 1 (25) • 5; 3 (75)	• 3; 2 (50) • 5; 2 (50)
I am confident today's session notes contain all necessary information	• 3; 1 (17) • 5; 5 (83)	• 5; 4 (100)	• 2; 1 (25) • 4; 1 (25) • 5; 2 (50)
I am confident today's session notes contain only relevant information	• 4; 1 (17) • 5; 5 (83)	• 5; 4 (100)	• 2; 1 (25) • 4; 2 (50) • 5; 1 (25)
I believe today's session notes will be valuable to others (parents, supervisors, behavior analyst)	• 4; 2 (33) • 5; 4 (67)	• 5; 4 (100)	• 2; 1 (25) • 3; 1 (25) • 4; 1 (25) • 5; 1 (25)

Post–Final Shift Survey Ratings and SUS Ratings

All 14 participants who used the DAT during multiple shifts completed the post–final shift survey. Table 3 shows the frequencies of responses to the first 7 statements.

The SUS score for the 14 participants was 81.79, which is quite high for an initial prototype; this score corresponds to *excellent* usability [32]. This score is also in the range where a product is likely to be recommended by users to other potential users [33].

The final survey also contained 11 open-ended questions: 4 (36%) were about the mobile app, 6 (55%) about the web application, and 1 (9%) asked if any initial concerns that participants had about using the DAT prototypes had been alleviated after using them during multiple shifts.

XSL•FO
RenderX

Table 3. Frequencies of responses to post–final shift survey non–System Usability Scale Likert scale items.

Statement	Site 1 (n=6), response; n (%)	Site 2 (n=4), response; n (%)	Site 3 (n=4), response; n (%)
I am confident that today's data sheets are accurate	• 4; 1 (17) • 5; 5 (83)	• 5; 4 (100)	• 3; 1 (25) • 4; 1 (25) • 5; 2 (50)
I found it easy to record behavior data for all clients today	• 5; 5 (100)	• 5; 4 (100)	• 3; 2 (50) • 4; 1 (25) • 5; 1 (25)
I believe today's behavior data will be valuable to others	• 4; 1 (17) • 5; 5 (83)	• 5; 4 (100)	• 3; 1 (25) • 4; 2 (50) • 5; 1 (25)
I found it easy to write session notes today	• 5; 5 (100)	• 5; 4 (100)	• 3; 2 (50) • 4; 1 (25) • 5; 1 (25)
I am confident today's session notes contain all necessary information	• 4; 1 (17) • 5; 5 (83)	• 5; 4 (100)	• 3; 1 (25) • 4; 1 (25) • 5; 2 (50)
I am confident today's session notes contain only relevant information	• 4; 2 (33) • 5; 4 (67)	• 5; 4 (100)	• 3; 2 (50) • 5; 2 (50)
I believe today's session notes will be valuable to others (parents, supervisors, behavior analyst)	• 4; 1 (17) • 5; 5 (83)	• 5; 4 (100)	• 3; 1 (25) • 4; 2 (50) • 5; 1 (25)

Thematic Analysis of Open-Ended Question Responses

Overview

One-third of the question responses (5/14, 36%) were coded by each of the 2 coders. Intercoder reliability was 92.56%. After coding was complete, the research assistants worked with the PI and identified the following 5 overarching themes in the survey response data.

Theme 1: Using the DAT Was a Positive Experience

Participants reported that the mobile app was easy to use and that it was helpful to be able to create notes in real time. Comments after using it for a single shift included "it's amazing" (Participant 4) and "I would use it on a daily basis" (Participant 12). The participants also provided predominantly positive feedback on the web application, such as "it's great," (Participant 2) "simple and easy," (Participant 7) and "organized and helpful" (Participant 6).

Theme 2: Using the DAT Made It Easier to Create End-of-Shift Reports

Survey responses also revealed that participants found that using the DAT facilitated writing end-of-shift reports. Their comments included "[The DAT] made it easy to keep track of everything throughout the day" (Participant 14) and "made [writing reports] a lot easier for me" (Participant 1). A participant wrote, "I had a great experience [writing reports] and I like [the DAT] very much" (Participant 7).

Theme 3: The DAT Helped Increase the Accuracy of End-of-Shift Reports

Participants indicated that using the DAT enabled them to create more accurate reports. One wrote, "It help [sic] to maintain accurate notes as the day goes along" (Participant 7); another reported that "It gives you the opportunity to keep the notes accurate" (Participant 9); and a third simply commented, "increases accuracy" (Participant 2).

Theme 4: Additional and Improved Features Are Desired

Although participants provided positive responses when asked about their experience using the DAT, they also noted that the tools could provide even greater benefit if they were enhanced. Suggestions included allowing users to access the web application to review and edit notes on their smartphone, improving the transcription accuracy; allowing users to store notes locally when smartphones are not connected to Wi-Fi; offering additional or more specific categories; and allowing users to copy time stamps when copying note text from the web application.

Theme 5: More or Better Training Would Help

The survey responses by 3 (21%) of the 14 participants suggested that they would have liked additional training. One of them noted that "it looks easy but when I'm by myself it's a different story" (Participant 12); the second wrote, "I wish I could get a better understanding" (Participant 11); and the third reported that "it would be easy if I get the hang of it" (Participant 12).

XSL•FO
RenderX

Discussion

Principal Findings

All 14 participants were able to use the mobile app without assistance during all shifts, but their ability to use the web application independently at the end of their first shift varied. All participants were able to use both the mobile app and the web application without help from the research team members during their second and third shifts. This result is consistent with an average SUS score of >80, indicating that the DAT prototypes are very easy to use. The hardest part about using the web application for most of the participants was typing in the URLs to access the private, secure notes review app. In future testing, aliases (shorter active links that take users to the same place as the longer URLs) should be set up to help avoid this problem. In addition, research team members should always help participants create a bookmark during training to make it easy to access the web application, a strategy that was not adopted until testing was conducted at site 3. All participants indicated that they found it easier to write reports at the end of their shift after using the DAT.

At site 1, all 6 participants were able to use the DAT prototypes for at least 2 successive shifts. These participants, whose ages were fairly well distributed across the range of 18 to 54 years, agreed or strongly agreed with 6 (86%) of the 7 survey statements after day 1. For the statement "I am confident today's session notes contain all necessary information," of the 6 participants, 1 (17%) selected *neither agree nor disagree*, and the remaining 5 (83%) all agreed or strongly agreed (Table 2). When given the same set of questions in the final survey, all 6 participants agreed or strongly agreed with all 7 statements (Table 3). This is consistent with the largely positive responses that these participants provided to the open-ended questions in the final survey. All indicated that they enjoyed using the DAT prototypes and would be interested in using future versions.

The site 2 participants were younger, with ages ranging from 18 to 34 years; therefore, they had less experience working with adults with IDDs than site 1 participants. Their responses to the Likert statements also indicated a high level of agreement; after day 1, all 4 participants strongly agreed with 6 (86%) of the 7 statements; 1 (25%) participant selected *neither agree nor disagree* for the statement "I found it easy to write session notes today," and the remaining 3 (75%) participants all strongly agreed with this statement. All 4 site 2 participants responded *strongly agree* to all 7 statements in the final survey. These participants also provided responses to the open-ended questions indicating that they enjoyed using the DAT prototypes and would be interested in using future versions. These 4 participants provided more suggestions for additional features than the participants from the other 2 sites.

Site 3 participants fell into 2 different age ranges. Of the 4 participants, 2 (50%) were aged between 18 and 34 years, and 2 (50%) were aged between 45 and 54 years. Compared to site 1 and site 2 participants, site 3 participants provided a wider range of responses to the 2 surveys. As can be seen in Table 2, after the first shift using the DAT prototypes, a participant disagreed with the statements "I am confident today's session notes contain all necessary information," "I am confident today's session notes contain only relevant information," and "I believe today's session notes will be valuable to others (parents, supervisors, behavior analyst)." Responses to the other 4 statements ranged from *neutral* to *strongly agree*. Responses to the final survey ranged from *neutral* to *strongly agree* for all 7 statements, indicating more variable sentiments among site 3 participants at the end of the pilot test than among site 1 and site 2 participants.

When we grouped the site 3 participants' responses based on their ages, we found that those in the age category of 45 to 54 years (2/4, 50%) generally reported a less positive experience using the DAT prototypes than the other participants (2/4, 50%). In fact, a review of the data logs from these 2 participants revealed that they only used the mobile app at the end of their shifts. It seemed that they used the DAT mostly as a transcription service. As the DAT prototypes used a freely available web-based transcription service which was far from perfect, and time stamps are not helpful when all notes are created at the end of shifts, these DSPs would have been better off using a transcription app on the computers they used to create their end-of-shift reports.

While most of the test participants (12/14, 86%) indicated that they valued the ability to collect data about their work in the moment, several of them (6/14, 43%) noted that their ability to benefit from this capability was reduced because the mobile app prototype requires Wi-Fi to save notes. Even with this limitation, the DSPs were very enthusiastic about the DAT prototypes; 1 participant asked whether she could keep using them after the pilot test, and another asked when they would be available for daily use. Future work includes enabling the mobile app to store data locally so that it can be used to collect data while DSPs and clients are out in the community, which is a key part of some DSPs' client interactions.

The average numbers of notes created per shift and canceled notes per shift was roughly the same for each site. The number of words per note averaged 6.5 at site 2 compared to approximately 10 at sites 1 and 3. To our surprise, audio notes had roughly the same number of words as text-based notes. While the number of words seemed low at first, this is consistent with the intent that note creation be quick, easy, and informal, helping jog memories during report creation. Meanwhile, the average number of notes that DSPs saved during their shifts increased for most DSPs across their shifts. We anticipated that the number of DSPs who created few notes initially would increase in subsequent shifts, and those who created many would create fewer as they learned which types of notes were most helpful. Although 1 DSP's notes fit this pattern, it seems that the other DSPs created more notes as they became more accustomed to using the mobile app during their shift. In a few cases, participants did not include any content; they just created a note with a client and category selected, indicating that just having a time associated with a client and category would be useful when writing reports. In addition, participants indicated that being able to sort notes by category can directly help with writing reports that have specific requirements, such as detailing all instances of medication administration or all food intake during a shift.

XSL•FO
RenderX

Overall, the results from the pilot test are promising, suggesting that DSPs would be willing to use mobile devices to enable in-the-moment data collection, provided that the collected data facilitate efficient generation of required end-of-shift documentation. Feedback from the test participants suggests that technology such as the DAT could help to increase efficiency, effectiveness, and job satisfaction among DSPs.

Comparison With Prior Literature

Many sources, some more than a decade old, have warned about, or described, a shortage of DSPs in the United States [7,34-39]. Factors such as heavy workload, onerous documentation requirements, and burnout contribute to a high turnover rate, which exacerbates stress on DSPs [22,40-49]. Some of this literature points out that technology could potentially help to reduce the time DSPs must spend on their other responsibilities, such as documentation, so that they can spend more time on direct support [7,38].

Technology has been successfully applied to increase documentation efficiency and decrease workload in special education and clinic- and home-based therapy programs for children with autism spectrum disorder [17,18,50,51]. By enabling users to quickly and easily record in-the-moment data, these technologies help improve the quality of documentation they generate [52]. However, there is relatively little work that explores how technology could be applied to support direct support workers, such as DSPs and home health aides (HHAs), who provide care to adults [53,54]. Most of the research that addresses having direct support workers leverage technology is qualitative and focuses on workers who support patients without IDDs [55-57]. One 2022 review paper identified only 1 study that created technology expressly intended to support the work of direct support workers [53]. This study, which also entailed creating a smartphone app to use to help create reports, was conducted in Japan nearly 2 decades ago. The authors who described this effort concluded that enabling direct support workers to use smartphones to create reports saves time and reduces costs [58]. Another 2022 review, which surveyed "the technological landscape" of direct support workers, noted that there is a paucity of evidence about how information and communication technologies can be used by these workers [57]. The authors of the review went on to assert that none of the existing technology-based interventions that could be used to facilitate home care were specifically designed to support the workflows of HHAs and concluded that "there is an urgent need for research that centers on the needs and perspectives of HHAs and using human-centered methods to engage HHAs in the design of technologies that truly support their essential caregiving work" [53]. The human-centered research reported here was specifically focused on developing technology for DSPs, but many of the design principles identified in Textbox 1 are also relevant to creating technologies to support the work of direct support workers who care for individuals without IDDs, such as HHAs.

However, inserting technology that is intended to facilitate sharing data among many care providers does not always increase documentation efficiency or decrease workload. Consider the migration from paper to electronic health records (EHRs), which significantly increased clinicians' workload [59-62]. EHRs have also been associated with a decrease in the amount of direct patient support time [63-65]. In contrast to EHRs, many of which grew out of billing systems [66], the DAT prototypes were developed by following an established user-centered design methodology [23] that ensured that our team considered DSPs' needs, goals, and constraints and obtained their feedback throughout our effort. As in the case of EHRs, a strong motivation behind creating the DAT prototypes helped to facilitate care for the individuals whose data are being logged electronically [67-70].

Multiple studies have pointed out that clinician burnout, which has been connected to the advent of EHRs, is correlated with decreased patient safety [71-74]. Similarly, DSP burnout and turnover are associated with poor health and quality of life for their clients [3,5]. Hence, technology such as the DAT prototypes, which are designed to reduce the burden of required documentation, improve job satisfaction, and reduce turnover among DSPs, can also help to improve the quality of life of adults with IDDs whom they support.

Limitations

Several limitations warrant mention. Our pilot test was limited to 14 DSPs working in 1 geographic region. However, different work settings, ranging from a private home with a single resident and group homes with 3 to 4 residents to a day program at a large agency, as well as the range of time that the DSPs who served as pilot testers had worked with adults with IDDs (<1 year to 20 years) increase the generalizability of our results. The pilot test length was also quite short, and we did not measure the amount of time spent generating documentation before the pilot test. These factors prevented us from assessing the impact of the DAT prototypes on documentation efficiency and quality. Nevertheless, the pilot test was long enough to establish feasibility.

Another limitation is that only 10 (71%) of the 14 pilot test volunteers were able to use the DAT prototypes for 3 successive shifts. Scheduling issues, often driven by staffing shortages that caused a staff member to be moved to a different location at the last minute, prevented several participants (4/14, 29%) from using the DAT as planned. In addition, we did not collect data about participants' familiarity and comfort with technology, although our observations while on site during the first shift provided insights about these factors. The fact that the DSPs were paid to participate is another limitation, although responses to the open-ended questions suggest that at least some of the participants were interested in continuing to use the tools when they would not be paid.

In addition, while the qualitative data from the test users was generally very positive, the study did not use any instruments that assess the likelihood of user adoption, such as the technology acceptance model [75]. Future work on technologies to support DSPs should be informed by the framework developed by Venkatesh et al [76] for understanding user acceptance of IT and might also consider the strategies outlined by Sebastian et al [77] for increasing user acceptance of novel technologies.

XSL•FO
RenderX

Implications

The results of the pilot test are promising, suggesting that upgraded versions of the DAT prototypes or similar technologies have the potential to reduce the burden of completing end-of-shift reports, while improving the quality of data produced by DSPs. Making it quick and simple to produce time-stamped in-the-moment notes facilitates logging more accurate and more detailed client data. Better data about adults with IDDs would enable family members, health providers, therapists, health providers, and behavior analysts to better support these adults. Long-term benefits of using the DAT could include (1) reducing DSP workload; (2) increasing the time DSPs spend interacting with adults with IDDs; (3) enabling DSPs to provide more consistent and appropriate support; (4) increasing DSP job satisfaction; (5) improving medical and behavioral support for adults with IDDs; and (6) providing a foundation for technology use that increases independence in adults with IDDs, thereby improving their quality of life. Digital documentation could also facilitate timely access to information about adults with IDDs for the diverse stakeholders who help support them.

However, there are risks associated with adopting technology-based data collection tools, including loss of data or an inability to log new data during technology failures. Moreover, to achieve the goal of capturing higher-quality data during work shifts, it will be necessary to allow users to customize the mobile app based on specific clients' support plans. This will add complexity, which could negatively impact user acceptance because many DSPs do not have a great deal of experience with technology. However, it is possible that supervisors or other employees at agencies that employ DSPs could learn to use an administrative portal to customize the mobile app based on clients' needs for support.

In any case, it will be important to obtain input from other stakeholders, such as behavioral supervisors, agency directors, and families, in future efforts, particularly as selection of suitable technology is likely the responsibility of the workspace and employer, and communication of data extends beyond DSPs. In the long term, technologies that support data collection will need to be integrated into existing report generation tools. Eventually, these technologies could even leverage artificial intelligence to create first drafts of DSPs' end-of-shift reports. If this sort of capability is developed, it will be very important to identify strategies for ensuring privacy and security and to consider the ethical implications of using artificial intelligence technologies [78,79].

Finally, while the DAT development effort focused on providing technology that supports DSPs, it is likely that many other types of direct workers could benefit from a similar platform (eg, HHAs, care workers, personal care assistants, certified nursing assistants, and nursing home assistants) [57,80-83]. This is significant because the United States is now facing a severe shortage across the entire direct support workforce [84,85] due in large part to the COVID-19 pandemic. As in the case with DSPs, the shortage in the direct support workforce is harmful to both these workers and those who rely upon them for support [86,87]. A report from the Centers for Disease Control and Prevention revealed that in 2016 a total of 61 million adults (approximately 1 in 4) living in the United States had a disability that impacts major life activities [88]. Many of these adults do, or will eventually, depend upon direct support workers for assistance with activities of daily living.

In summary, future efforts should not only increase the capabilities and robustness of the initial DAT prototypes and consider the needs of family members, medical providers, behavior analysts, and others who would benefit from timely accurate data about adults with IDDs but also explore how these tools could be adapted to meet the needs of other types of direct support workers by eliciting information and feedback from these workers.

Conclusions

DSPs play a critical role in the care of adults with IDDs. Technology can help mitigate the high turnover rates, poor job satisfaction, and the burden of necessary data collection and documentation that negatively impact DSPs' ability to care for these adults. The user-centered research effort reviewed here produced proof-of-concept prototypes of tools intended to improve the effectiveness and job satisfaction of DSPs. The results of the pilot test indicate that these tools are likely to provide the intended benefits to DSPs and thus have the potential to help improve the quality of life of clients served by DSPs.

Future research should include testing more robust feature-enhanced versions of the DAT over longer periods in even more diverse settings where DSPs provide support to adults with IDDs. Additional work should also include identifying or developing an instrument to reliably assess report quality and time-motion studies of DSPs before and during longer trials to help quantify how much time DSPs spend generating required documentation. One time-motion study of physicians working on hospital wards found that the physicians believed that they were spending more time on documentation and other administrative tasks than they actually were [89].

In any case, the work reported here, despite its limitations, provides valuable insights into how technology could benefit DSPs and the people they support. Feedback from DSPs indicates that the highest-priority feature enhancement for the DAT prototypes is enabling the mobile app to store data locally to support in-the-moment data collection without Wi-Fi connectivity. Several other enhancements, such as shared task lists, were identified as part of the initial user needs analysis activities performed at the start of this effort [16]. In addition to adding some of these capabilities, future work should identify the needs and constraints of the service providers who employ DSPs to identify barriers to adopting data collection and documentation technologies, such as costs, adaptability in small operations, the need to protect confidentiality, minimizing potential technology damage, and preventing data loss. Such work could help enable future versions of the DAT to supply all caregivers and service providers with the information necessary for better overall service, outcomes, and quality of life for adults with IDDs.

Finally, this line of research needs to be expanded because it could have a profound impact on the health and welfare of

XSL•FO
RenderX

several other adults beyond those with IDDs who are supported by direct support workers: older adults, individuals with physical disabilities, and individuals with severe mental illness. Furthermore, the direct support workers themselves could also benefit from technologies such as DATs that enable quick and easy in-the-moment data collection and facilitate end-of-shift reporting.

Acknowledgments

Funding for this work was provided by a grant from the New Jersey Department of Health Governor's Council for Medical Research and Treatment of Autism (CAUT2020APL016). The authors would like to thank the direct support professionals who participated in these studies for sharing their valuable experiences and perspectives, the leadership of the 3 organizations where we conducted the pilot tests, and Adam Jonas and Morgan Murphy for helping with the thematic analysis of the responses to the open-ended survey questions.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Initial designs for the digital assistant tool mobile app.
[PDF File (Adobe PDF File), 938 KB - humanfactors_v11i1e51612_app1.pdf]

Multimedia Appendix 2
Demographic survey filled out by all pilot test participants.
[DOCX File , 21 KB - humanfactors_v11i1e51612_app2.docx]

Multimedia Appendix 3
Post–shift 1 survey filled out after using the digital assistant tool prototypes for a single shift.
[DOCX File , 16 KB - humanfactors_v11i1e51612_app3.docx]

Multimedia Appendix 4
Final survey filled out after participants had used the digital assistant tool prototypes for 2 or 3 shifts.
[DOCX File , 16 KB - humanfactors_v11i1e51612_app4.docx]

References

1. Larson SA, Neidorf J, Pettingell SL, Sowers M. In-home and residential long-term supports and services for persons with intellectual or developmental disabilities: status and trends through 2019. Minneapolis: University of Minnesota, Institute on Community Integration. 2023. URL: https://ici.umn.edu/products/AiMjLBqCS8mSwjLmIPPvKg [accessed 2024-04-03]
2. Mlinac ME, Feng MC. Assessment of activities of daily living, self-care, and independence. Arch Clin Neuropsychol 2016 Sep 29;31(6):506-516. [doi: 10.1093/arclin/acw049] [Medline: 27475282]
3. Friedman C. Direct support professionals and quality of life of people with intellectual and developmental disabilities. Intellect Dev Disabil 2018 Aug;56(4):234-250. [doi: 10.1352/1934-9556-56.5.234] [Medline: 30024847]
4. Friedman C. The impact of continuity and security on quality of life. Intellect Dev Disabil 2022 Apr 01;60(2):101-112. [doi: 10.1352/1934-9556-60.2.101] [Medline: 35297986]
5. Friedman C. The impact of direct support professional turnover on the health and safety of people with intellectual and developmental disabilities. Inclusion. 2021 2021;9(1):63-73 [FREE Full text] [doi: 10.1352/2326-6988-9.1.63]
6. Personal outcome measures: measuring personal quality of life. The Council on Quality and Leadership. 2017. URL: https://www.c-q-l.org/wp-content/uploads/2020/03/2017-CQL-POM-Manual-Adults.pdf [accessed 2024-04-03]
7. Macbeth JM. Commitment, capacity and culture: solutions for the direct support workforce crisis. The National Alliance for Direct Support Professionals. 2011. URL: https://tinyurl.com/7w8v9k35 [accessed 2024-04-03]
8. Claes C, Van Hove G, Vandevelde S, van Loon J, Schalock R. The influence of supports strategies, environmental factors, and client characteristics on quality of life-related personal outcomes. Res Dev Disabil 2012 Jan;33(1):96-103. [doi: 10.1016/j.ridd.2011.08.024] [Medline: 22093653]
9. Felce D, Lowe K, Jones E. Staff activity in supported housing services. Res Intellect Disabil 2002 Dec 11;15(4):388-403. [doi: 10.1046/j.1468-3148.2002.00130.x]
10. Felce D, Perry J. Living with support in the community. In: Odom SL, Horner RH, Snell ME, Blacher JB, editors. Handbook of Developmental Disabilities. Oxfordshire, United Kingdom: Routledge; 2007:410-428.

XSL•FO
RenderX

11. McConkey R, Abbott S, Walsh PN, Linehan C, Emerson E. Variations in the social inclusion of people with intellectual disabilities in supported living schemes and residential settings. J Intellect Disabil Res 2007 Mar 18;51(Pt 3):207-217. [doi: 10.1111/j.1365-2788.2006.00858.x] [Medline: 17300416]
12. Li S. Direct care personnel recruitment, retention, and orientation. Community-University Institute for Social Research. 2004. URL: https://tinyurl.com/4a7awczz [accessed 2024-04-03]
13. Brandt J. Report to the president 2017: America's direct support workforce crisis. President's Committee for People with Intellectual Disabilities (PCPID). 2017. URL: https://tinyurl.com/mry7afc8 [accessed 2024-04-03]
14. Hall SA, Anderson LL, Pettingell SL, Zhang A, Bershadsky J, Hewitt A, et al. Direct support professional and frontline supervisor perspectives on work-life in a pandemic. Inclusion (Wash) 2022;10(4):314-326 [FREE Full text] [doi: 10.1352/2326-6988-10.4.314] [Medline: 38528916]
15. Friedman C. The COVID-19 pandemic and quality of life outcomes of people with intellectual and developmental disabilities. Disabil Health J 2021 Oct;14(4):101117 [FREE Full text] [doi: 10.1016/j.dhjo.2021.101117] [Medline: 34144897]
16. Simmons CA, Moretti AE, Lobo AF, Tremoulet PD. Direct support professionals' perspectives on using technology to help support adults with autism spectrum disorder: mixed methods study. JMIR Form Res 2023 Apr 25;7:e40722 [FREE Full text] [doi: 10.2196/40722] [Medline: 37097738]
17. Hayes GR, Gardere LM, Abowd GD, Truong KN. CareLog: a selective archiving tool for behavior management in schools. In: Proceedings of the 2008 SIGCHI Conference on Human Factors in Computing Systems. 2008 Presented at: CHI '08; April 5-10, 2008; Florence, Italy p. 685-694 URL: https://dl.acm.org/doi/10.1145/1357054.1357164 [doi: 10.1145/1357054.1357164]
18. Bangerter A, Manyakov NV, Lewin D, Boice M, Skalkin A, Jagannatha S, et al. Caregiver daily reporting of symptoms in autism spectrum disorder: observational study using web and mobile apps. JMIR Ment Health 2019 Mar 26;6(3):e11365 [FREE Full text] [doi: 10.2196/11365] [Medline: 30912762]
19. Sleeper JD, LeBlanc LA, Mueller J, Valentino AL, Fazzio D, Raetz PB. The effects of electronic data collection on the percentage of current clinician graphs and organizational return on investment. J Organ Behav Manage 2017 Jan 10;37(1):83-95. [doi: 10.1080/01608061.2016.1267065]
20. Kroezen M, Schäfer W, Sermeus W, Hansen J, Batenburg R. Healthcare assistants in EU member states: an overview. Health Policy 2018 Oct;122(10):1109-1117. [doi: 10.1016/j.healthpol.2018.07.004] [Medline: 30033205]
21. Desa U. World population prospects 2019: highlights. United Nations Department for Economic and Social Affairs. 2019. URL: https://population.un.org/wpp/Publications/Files/WPP2019_10KeyFindings.pdf [accessed 2024-04-03]
22. Lakin KC, Burke MM. Looking forward: research to respond to a rapidly aging population. Res Pract Persons Severe Disabl 2019 Nov 01;44(4):280-292. [doi: 10.1177/1540796919882356]
23. Craven PL, Tremoulet PD, Regli SH. Incorporating human systems engineering in advanced military technology development. In: McNeese MD, Forster PK, editors. Cognitive Systems Engineering: An Integrative Living Laboratory Framework. Boca Raton, FL: CRC Press; 2017:341-359.
24. Isaac S, Michael WB. Handbook in Research and Evaluation: A Collection of Principles, Methods, and Strategies Useful in the Planning, Design, and Evaluation of Studies in Education and the Behavioral Sciences. 3rd edition. New York, NY: Edits Publications; 1995.
25. Hill R. What sample size is "enough" in internet survey research. Interpers Compute Technol 1998;6(3-4):1-12 [FREE Full text]
26. Julious SA. Sample size of 12 per group rule of thumb for a pilot study. Pharm Stat 2005 Nov 24;4(4):287-291. [doi: 10.1002/pst.185]
27. van Belle G. Statistical Rules of Thumb. Hoboken, NJ: John Wiley & Sons; 2011.
28. Brooke J. SUS: a retrospective. J Usability Stud 2013;8(2):29-40 [FREE Full text]
29. Brooke J. SUS: a 'quick and dirty' usability scale. In: Jordan PW, Thomas B, McClelland IL, Weerdmeester B, editors. Usability Evaluation In Industry. Boca Raton, FL: CRC Press; 1996:4-7.
30. Braun V, Clarke V. Using thematic analysis in psychology. Qual Res Psychol 2006 Jan;3(2):77-101. [doi: 10.1191/1478088706qp063oa]
31. Maguire M, Delahunt B. Doing a thematic analysis: a practical, step-by-step guide for learning and teaching scholars. Ir J High Educ 2017;9(3) [FREE Full text]
32. Bangor A, Kortum P, Miller J. Determining what individual SUS scores mean: adding an adjective rating scale. J Usability Stud 2009;4(3):114-123 [FREE Full text]
33. Sauro J. Does better usability increase customer loyalty? MeasuringU. 2010. URL: https://measuringu.com/usability-loyalty/ [accessed 2024-04-03]
34. Swanson-Aprill L, Luz C, Travis A, Hunt J, Wamsley S. Policy brief direct care workforce shortage. Michigan Aging and Adult Services Agency. 2019. URL: https://tinyurl.com/2s3n949n [accessed 2024-04-02]
35. Taylor SJ. The direct support workforce crisis: can unions help resolve this? Center on Human Policy. 2008. URL: https://eric.ed.gov/?id=ED503912 [accessed 2024-04-02]
36. Bradley VJ. How COVID-19 may change the world of services to people with intellectual and developmental disabilities. Intellect Dev Disabil 2020 Oct 01;58(5):355-360. [doi: 10.1352/1934-9556-58.5.355] [Medline: 33032314]

XSL•FO
RenderX

37. Embregts PJ, Tournier T, Frielink N. Experiences and needs of direct support staff working with people with intellectual disabilities during the COVID-19 pandemic: a thematic analysis. J Appl Res Intellect Disabil 2021 Mar 21;34(2):480-490 [FREE Full text] [doi: 10.1111/jar.12812] [Medline: 32954585]
38. Position statement on the direct support professional workforce. New Jersey Council on Developmental Disabilities. 2019. URL: https://njcdd.org/wp-content/uploads/DSP-NJCDD-Position-Statement.pdf [accessed 2024-04-03]
39. Laws CB. The direct support professional workforce crisis: challenges, state approaches, and opportunities for Georgia. The University of Georgia. 2019. URL: https://tinyurl.com/3pt3c3s8 [accessed 2024-04-03]
40. Pettingell SL, Houseworth J, Tichá R, Kramme JE, Hewitt AS. Incentives, wages, and retention among direct support professionals: national core indicators staff stability survey. Intellect Dev Disabil 2022 Apr 01;60(2):113-127 [FREE Full text] [doi: 10.1352/1934-9556-60.2.113] [Medline: 35297985]
41. Thornton CR. Direct support staff retention and turnover in the field of applied behavior analysis: a national survey. Temple University Graduate Board. 2018. URL: https://tinyurl.com/4zbdkpp6 [accessed 2024-04-05]
42. Kazemi E, Shapiro M, Kavner A. Predictors of intention to turnover in behavior technicians working with individuals with autism spectrum disorder. Res Autism Spectr Disord 2015 Sep;17:106-115. [doi: 10.1016/j.rasd.2015.06.012]
43. Bogenschutz M, Nord D, Hewitt A. Competency-based training and worker turnover in community supports for people with IDD: results from a group randomized controlled study. Intellect Dev Disabil 2015 Jun;53(3):182-195. [doi: 10.1352/1934-9556-53.3.182] [Medline: 26107852]
44. Bossink LW, van der Putten AA, Vlaskamp C. Understanding low levels of physical activity in people with intellectual disabilities: a systematic review to identify barriers and facilitators. Res Dev Disabil 2017 Sep;68:95-110. [doi: 10.1016/j.ridd.2017.06.008] [Medline: 28750208]
45. Gray-Stanley JA, Muramatsu N, Heller T, Hughes S, Johnson TP, Ramirez-Valles J. Work stress and depression among direct support professionals: the role of work support and locus of control. J Intellect Disabil Res 2010 Aug 15;54(8):749-761 [FREE Full text] [doi: 10.1111/j.1365-2788.2010.01303.x] [Medline: 20633203]
46. Couderc S, Cousson-Gélie F, Pernon E, Porro B, Miot S, Baghdadli A. Burnout among direct support workers of adults with autism spectrum disorder and intellectual disability. Scand J Caring Sci 2023 Mar;37(1):131-140. [doi: 10.1111/scs.13008] [Medline: 34075606]
47. Gray-Stanley JA, Muramatsu N. Work stress, burnout, and social and personal resources among direct care workers. Res Dev Disabil 2011 May;32(3):1065-1074 [FREE Full text] [doi: 10.1016/j.ridd.2011.01.025] [Medline: 21316918]
48. Mutkins E, Brown R, Thorsteinsson E. Stress, depression, workplace and social supports and burnout in intellectual disability support staff. J Intellect Disabil Res 2011 May;55(5):500-510. [doi: 10.1111/j.1365-2788.2011.01406.x] [Medline: 21418365]
49. Ryan C, Bergin M, Wells JS. Work-related stress and well-being of direct care workers in intellectual disability services: a scoping review of the literature. Int J Dev Disabil 2019 Mar 27;67(1):1-22 [FREE Full text] [doi: 10.1080/20473869.2019.1582907] [Medline: 34141395]
50. Siyam N. Factors impacting special education teachers' acceptance and actual use of technology. Educ Inf Technol 2019 Jan 8;24(3):2035-2057. [doi: 10.1007/s10639-018-09859-y]
51. Bak MY, Plavnick JB, Dueñas AD, Brodhead MT, Avendaño SM, Wawrzonek AJ, et al. The use of automated data collection in applied behavior analytic research: a systematic review. Behav Anal Res Pract 2021 Nov;21(4):376-405. [doi: 10.1037/bar0000228]
52. Nadkar RS. Increasing accuracy of data collection by direct support professionals with task clarification and performance feedback. Temple University Libraries. 2018. URL: https://scholarshare.temple.edu/handle/20.500.12613/1990 [accessed 2024-04-03]
53. Kuo EF, Cho J, Olaye I, Delgado D, Dell N, Sterling MR. Understanding the technological landscape of home health aides: scoping literature review and a landscape analysis of existing mHealth apps. J Med Internet Res 2022 Nov 11;24(11):e39997 [FREE Full text] [doi: 10.2196/39997] [Medline: 36176033]
54. Bartle V, Albright L, Dell N. "This machine is for the aides": tailoring voice assistant design to home health care work. In: Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems. 2023 Presented at: CHI '23; April 23-28, 2023; Hamburg, Germany p. 1-9 URL: https://dl.acm.org/doi/abs/10.1145/3544548.3581118 [doi: 10.1145/3544548.3581118]
55. Franzosa E, Judon KM, Gottesman EM, Koufacos NS, Runels T, Augustine M, et al. Home health aides' increased role in supporting older veterans and primary healthcare teams during COVID-19: a qualitative analysis. J Gen Intern Med 2022 Jun 22;37(8):1830-1837 [FREE Full text] [doi: 10.1007/s11606-021-07271-w] [Medline: 35319082]
56. Islam SM, Halooq A, Dening J, Uddin R, Laranjo L, Chow CK, et al. Healthcare providers' perspectives on using smart home systems to improve self-management and care in people with heart failure: a qualitative study. Int J Med Inform 2022 Nov;167:104837. [doi: 10.1016/j.ijmedinf.2022.104837] [Medline: 36126353]
57. Perez H, Neubauer N, Marshall S, Philip S, Miguel-Cruz A, Liu L. Barriers and benefits of information communication technologies used by health care aides. Appl Clin Inform 2022 Jan 09;13(1):270-286 [FREE Full text] [doi: 10.1055/s-0042-1743238] [Medline: 35263800]
58. Ogawa H, Yonezawa Y, Maki H, Sato H, Hahn AW, Caldwell WM. A Java mobile phone-based "home helper" care report creation support system. Biomed Sci Instrum 2004;40:76-79. [Medline: 15133938]

XSL•FO
RenderX

59. Lou SS, Lew D, Harford DR, Lu C, Evanoff BA, Duncan JG, et al. Temporal associations between EHR-derived workload, burnout, and errors: a prospective cohort study. J Gen Intern Med 2022 Jul 16;37(9):2165-2172 [FREE Full text] [doi: 10.1007/s11606-022-07620-3] [Medline: 35710654]
60. Shan Y, Shang J, Yan Y, Ye X. Workflow interruption and nurses' mental workload in electronic health record tasks: an observational study. BMC Nurs 2023 Mar 09;22(1):63 [FREE Full text] [doi: 10.1186/s12912-023-01209-9] [Medline: 36890555]
61. Arndt BG, Beasley JW, Watkinson MD, Temte JL, Tuan WJ, Sinsky CA, et al. Tethered to the EHR: primary care physician workload assessment using EHR event log data and time-motion observations. Ann Fam Med 2017 Sep 11;15(5):419-426 [FREE Full text] [doi: 10.1370/afm.2121] [Medline: 28893811]
62. Gregory ME, Russo E, Singh H. Electronic health record alert-related workload as a predictor of burnout in primary care providers. Appl Clin Inform 2017 Dec 20;08(03):686-697. [doi: 10.4338/aci-2017-01-ra-0003]
63. Baumann LA, Baker J, Elshaug AG. The impact of electronic health record systems on clinical documentation times: a systematic review. Health Policy 2018 Aug;122(8):827-836. [doi: 10.1016/j.healthpol.2018.05.014] [Medline: 29895467]
64. Joukes E, Abu-Hanna A, Cornet R, de Keizer N. Time spent on dedicated patient care and documentation tasks before and after the introduction of a structured and standardized electronic health record. Appl Clin Inform 2018 Jan 17;9(1):46-53 [FREE Full text] [doi: 10.1055/s-0037-1615747] [Medline: 29342479]
65. Asan O, D Smith P, Montague E. More screen time, less face time - implications for EHR design. J Eval Clin Pract 2014 Dec 19;20(6):896-901 [FREE Full text] [doi: 10.1111/jep.12182] [Medline: 24835678]
66. Kim E, Rubinstein SM, Nead KT, Wojcieszynski AP, Gabriel PE, Warner JL. The evolving use of electronic health records (EHR) for research. Semin Radiat Oncol 2019 Oct;29(4):354-361. [doi: 10.1016/j.semradonc.2019.05.010] [Medline: 31472738]
67. Transforming health care through information technology. President's Information Technology Advisory Committee Panel on Transforming Health Care. 2006. URL: https://www.nitrd.gov/pubs/pitac/pitac-hc-9feb01.pdf [accessed 2024-04-03]
68. Evans RS. Electronic health records: then, now, and in the future. Yearb Med Inform 2018 Mar 06;25(S 01):S48-S61. [doi: 10.15265/iys-2016-s006]
69. Kohli R, Tan SS. Electronic health records: how can IS researchers contribute to transforming healthcare? MIS Q 2016;40(3):553-574 [FREE Full text]
70. Burde H. The HITECH act: an overview. AMA J Ethics 2011;13(3):172-175 [FREE Full text]
71. Nijor S, Rallis G, Lad N, Gokcen E. Patient safety issues from information overload in electronic medical records. J Patient Saf 2022 Sep 01;18(6):e999-1003 [FREE Full text] [doi: 10.1097/PTS.0000000000001002] [Medline: 35985047]
72. Aldosari B. Patients' safety in the era of EMR/EHR automation. Inform Med Unlocked 2017;9:230-233. [doi: 10.1016/j.imu.2017.10.001]
73. Middleton B, Bloomrosen M, Dente MA, Hashmat B, Koppel R, Overhage JM, American Medical Informatics Association. Enhancing patient safety and quality of care by improving the usability of electronic health record systems: recommendations from AMIA. J Am Med Inform Assoc 2013 Jun 01;20(e1):e2-e8 [FREE Full text] [doi: 10.1136/amiajnl-2012-001458] [Medline: 23355463]
74. Meeks DW, Smith MW, Taylor L, Sittig DF, Scott JM, Singh H. An analysis of electronic health record-related patient safety concerns. J Am Med Inform Assoc 2014 Nov 01;21(6):1053-1059 [FREE Full text] [doi: 10.1136/amiajnl-2013-002578] [Medline: 24951796]
75. Davis FD. Technology acceptance model: TAM. In: Al-Aufi AS, Al-Suqri MN, editors. Information Seeking Behavior and Technology Adoption: Theories and Trends. Hershey, PA: IGI Global; 1989:205-219.
76. Venkatesh V, Morris MG, Davis GB, Davis FD. User acceptance of information technology: toward a unified view. MIS Q 2003;27(3):425. [doi: 10.2307/30036540]
77. Sebastian G, George A, Jackson Jr G. Persuading patients using rhetoric to improve artificial intelligence adoption: experimental study. J Med Internet Res 2023 Mar 13;25:e41430 [FREE Full text] [doi: 10.2196/41430] [Medline: 36912869]
78. Sebastian G. Privacy and data protection in ChatGPT and other AI chatbots: strategies for securing user information. SSRN J 2023 [FREE Full text] [doi: 10.2139/ssrn.4454761]
79. Sebastian G. Exploring ethical implications of ChatGPT and other AI chatbots and regulation of disinformation propagation. SSRN J 2023. [doi: 10.2139/ssrn.4461801]
80. Ming J, Kuo E, Go K, Tseng E, Kallas J, Vashistha A, et al. "I go beyond and beyond" examining the invisible work of home health aides. Proc ACM Hum Comput Interact 2023 Apr 16;7(CSCW1):1-21. [doi: 10.1145/3579492]
81. Saari M, Patterson E, Killackey T, Raffaghello J, Rowe A, Tourangeau AE. Home-based care: barriers and facilitators to expanded personal support worker roles in Ontario, Canada. Home Health Care Serv Q 2017 Nov 08;36(3-4):127-144. [doi: 10.1080/01621424.2017.1393482] [Medline: 29048246]
82. King S, Liu L, Stroulia E, Nikolaidis I. Using simulations to integrate technology into health care aides' workflow. Int J Adv Corp Learn 2013 Jul 23;6(2):28. [doi: 10.3991/ijac.v6i2.2982]
83. Sterling MR, Dell N, Tseng E, Okeke F, Cho J, Piantella B, et al. Home care workers caring for adults with heart failure need better access to training and technology: a role for implementation science. J Clin Transl Sci 2020 Apr 06;4(3):224-228 [FREE Full text] [doi: 10.1017/cts.2020.36] [Medline: 32695493]

XSL•FO
RenderX

84. Kreider AR, Werner RM. The home care workforce has not kept pace with growth in home and community-based services. Health Aff (Millwood) 2023 May 01;42(5):650-657. [doi: 10.1377/hlthaff.2022.01351] [Medline: 37075251]
85. Jumabhoy S, Jung HY, Yu J. Characterizing the direct care health workforce in the United States, 2010-2019. J Am Geriatr Soc 2022 Feb 22;70(2):512-521. [doi: 10.1111/jgs.17519] [Medline: 34687042]
86. Smallwood N, Harrex W, Rees M, Willis K, Bennett CM. COVID-19 infection and the broader impacts of the pandemic on healthcare workers. Respirology 2022 Jun 19;27(6):411-426 [FREE Full text] [doi: 10.1111/resp.14208] [Medline: 35048469]
87. Harootunian L, Buffett A, O'Gara B, Perry K, Serafini MW, Hoagland GW. A multipronged approach to alleviating the direct care workforce shortage. Health Aff forefront 2023 [FREE Full text] [doi: 10.1377/forefront.20230608.815571]
88. Okoro CA, Hollis ND, Cyrus AC, Griffin-Blake S. Prevalence of disabilities and health care access by disability status and type among adults - United States, 2016. MMWR Morb Mortal Wkly Rep 2018 Aug 17;67(32):882-887 [FREE Full text] [doi: 10.15585/mmwr.mm6732a3] [Medline: 30114005]
89. Westbrook JI, Ampt A, Kearney L, Rob MI. All in a day's work: an observational study to quantify how and with whom doctors on hospital wards spend their time. Med J Aust 2008 May 05;188(9):506-509. [doi: 10.5694/j.1326-5377.2008.tb01762.x] [Medline: 18459920]

Abbreviations

DAT: digital assistant tool
DSP: direct support professional
EHR: electronic health record
HHA: home health aide
IDD: intellectual and developmental disability
IDEAS: Interaction Design and Engineering for Advanced Systems
PI: principal investigator
SUS: System Usability Scale

Edited by A Parush; submitted 05.08.23; peer-reviewed by G Sebastian, A AL-Asadi; comments to author 14.01.24; revised version received 07.03.24; accepted 16.03.24; published 25.04.24.

Please cite as:
Tremoulet PD, Lobo AF, Simmons CA, Baliga G, Brady M
Assessing the Usability and Feasibility of Digital Assistant Tools for Direct Support Professionals: Participatory Design and Pilot-Testing
JMIR Hum Factors 2024;11:e51612
URL: https://humanfactors.jmir.org/2024/1/e51612
doi:10.2196/51612
PMID:38662420

©Patrice D Tremoulet, Andrea F Lobo, Christina A Simmons, Ganesh Baliga, Matthew Brady. Originally published in JMIR Human Factors (https://humanfactors.jmir.org), 25.04.2024. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on https://humanfactors.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Assessing the Feasibility and Preliminary Effects of a Web-Based Self-Management Program for Chronic Noncancer Pain: Mixed Methods Study

Pascale Marier-Deschenes[1,2], PhD; Anne Marie Pinard[1,2,3], MEd, MD; Laura Jalbert[4], BA, MSc; Annie LeBlanc[1,4], PhD

[1]Laval University, Medicine Faculty, Québec, QC, Canada
[2]CIRRIS, Centre interdisciplinaire de recherche en réadaptation et intégration sociale, Québec, QC, Canada
[3]CHU de Québec, Université Laval, Québec, QC, Canada
[4]VITAM, Centre de recherche en santé durable, Québec, QC, Canada

Corresponding Author:
Pascale Marier-Deschenes, PhD
CIRRIS, Centre interdisciplinaire de recherche en réadaptation et intégration sociale
525, boul. Wilfrid-Hamel Québec
Québec, QC, G1M 2S8
Canada
Phone: 1 418 649 3735
Email: pascale.marier-deschenes.1@ulaval.ca

Abstract

Background: In Canada, adults with chronic noncancer pain face a persistent insufficiency of publicly funded resources, with the gold standard multidisciplinary pain treatment facilities unable to meet the high clinical demand. Web-based self-management programs cost-effectively increase access to pain management and can improve several aspects of physical and emotional functioning. Aiming to meet the demand for accessible, fully automated resources for individuals with chronic noncancer pain, we developed a French web- and evidence-based self-management program, *Agir pour moi* (APM). This program includes pain education and strategies to reduce stress, practice mindfulness, apply pacing, engage in physical activity, identify and manage thinking traps, sleep better, adapt diet, and sustain behavior change.

Objective: This study aims to assess the APM self-management program's feasibility, acceptability, and preliminary effects in adults awaiting specialized services from a center of expertise in chronic pain management.

Methods: We conducted a mixed methods study with an explanatory sequential design, including a web-based 1-arm trial and qualitative semistructured interviews. We present the results from both phases through integrative tables called joint displays.

Results: Response rates were 70% (44/63) at postintervention and 56% (35/63) at 3-month follow-up among the 63 consenting participants who provided self-assessed information at baseline. In total, 46% (29/63) of the participants completed the program. We interviewed 24% (15/63) of the participants. The interview's first theme revolved around the overall acceptance, user-friendliness, and engaging nature of the program. The second theme emphasized the differentiation between microlevel and macrolevel engagements. The third theme delved into the diverse effects observed, potentially influenced by the macrolevel engagements. Participants highlighted the features that impacted their self-efficacy and the adoption of self-management strategies. We observed indications of improvement in self-efficacy, pain intensity, pain interference, depression, and catastrophizing. Interviewees described these and various other effects as potentially influenced by macrolevel engagement through behavioral change.

Conclusions: These findings provided preliminary evidence that the APM self-management program and research methods are feasible. However, some participants expressed the need for at least phone reminders and minimal support from a professional available to answer questions over the first few weeks of the program to engage. Recruitment strategies of a future randomized controlled trial should focus on attracting a broader representation of individuals with chronic pain in terms of gender and ethnicity.

Trial Registration: ClinicalTrials.gov NCT05319652; https://clinicaltrials.gov/study/NCT05319652

(JMIR Hum Factors 2024;11:e50747) doi:10.2196/50747

XSL•FO
RenderX

KEYWORDS

persistent pain; eHealth; self-paced intervention; web-based program; evidence based; web based; self-management; pain; chronic pain; mixed methods study; pain treatment; pain education

Introduction

Background

The prevalence of chronic pain, including chronic cancer pain in adults, is estimated to be between 18% and 21%, with severe repercussions for all aspects of the lives of those affected, their families, and society [1-3]. Chronic pain affects patient-perceived health status and psychological functioning; decreases energy levels; and hinders engagement with physical, emotional, cognitive, and social activities [3-6]. These impacts can strain familial and social relationships and affect work performance [7]. Living with chronic pain often involves increased medical expenditures and detrimentally affects one's financial well-being [3,6]. In addition, the wait for services is not without added consequences to these repercussions, with long wait time (12-30 months) being associated with further deterioration in pain-related interference, psychological distress, and pain acceptance [3,8-10].

In Canada, adults with chronic noncancer pain face a persistent insufficiency of publicly funded resources, with the gold standard multidisciplinary pain treatment facilities being unable to meet the high clinical demand [1,2,11]. Since 2019, the Canadian Task Force has reaffirmed the necessity to implement equitable and innovative ways to deliver health interventions in a timely manner in the public network [12,13]. Web-based self-management programs that include exercise, sleep hygiene, pacing, and a healthy lifestyle are endorsed as part of the therapeutic considerations and recommendations for chronic noncancer pain management [14]. These programs have shown an impact on patients' pain intensity, pain interference [15,16], anxiety [15,17], depression [17,18], stress [18], catastrophizing, and self-efficacy [19].

The lack of accessible and reliable unguided web-based self-management programs tailored to French-speaking individuals with chronic noncancer pain is a significant yet solvable health services gap. Over the years, individuals with lived experience, organizations, and researchers have stressed the relevance and importance of actively involving patient partners in the health intervention development process [20-25]. Therefore, a novel French web-based self-management program for chronic noncancer pain developed in collaboration with individuals with lived experience could meet the specific needs of French-speaking individuals [26].

Objectives

This study aims to (1) assess the feasibility and acceptability of the *Agir pour moi* (APM) self-management program and trial procedures and (2) explore preliminary outcomes in individuals living with chronic noncancer pain.

Methods

Study Design

We conducted a mixed methods sequential explanatory study consisting of a single-arm, pre- and postintervention trial, followed by qualitative, semistructured interviews with adults experiencing chronic noncancer pain and awaiting services from a center of expertise in chronic pain management [27-29]. We registered the trial at ClinicalTrials.gov (NCT05319652) and followed the CONSORT-EHEALTH (Consolidated Standards of Reporting Trials of Electronic and Mobile Health Applications and Online Telehealth) and guidelines for reporting nonrandomised pilot and feasibility studies [30,31].

Ethical Considerations

The University Hospital Centre (CHU) de Québec-Université Laval Research Ethics Board approved the study (#2023-6312).

Knowledge Users' Involvement

A total of 7 individuals with lived experience of chronic pain, 5 health care professionals with experience and expertise in chronic pain management, 3 medical students, and 1 graphic designer contributed to the program's codevelopment. We engaged in web-based, phone, or email conversations over a period of 1.5 years. All knowledge users could contribute to various aspects of program development, including its identity (eg, colors, logo, and name), structure (eg, lesson sequence, content organization, and navigation), content (eg, self-management strategies, theoretical content, and testimonials), and learning modalities and behavior change techniques (eg, personal plans, reflective activities, and interactive scenarios). These knowledge users were not further involved across the duration of the trial, but individuals with lived experience initially guided the team toward reducing the questionnaire burden to a minimum for the participants.

Study Setting, Participants, and Recruitment

We recruited participants from the center of expertise in chronic pain management waitlist at the CHU de Québec-Université Laval. This center provides superspecialized services intended for complex chronic pain cases requiring a technical platform and multidisciplinary team. Most individuals referred to such centers experience significant impairments, including high pain levels interfering with their daily life, moderate to extremely severe depression, and pain-related sleep disturbance. Most of them take prescription analgesic medication and have already consulted different types of health care professionals [2].

Using the center's assigned priority level, between June and August 2022, we sent 500 invitation letters to adults (aged >18 years) with chronic noncancer pain (for >3 months) unlikely to receive services within the next 6 months. Interested individuals were to email us to set an eligibility interview, confirming that they understood French, had access to a computer and high-speed internet, had not started a new treatment for pain

XSL•FO
RenderX

within the last 1 month and agreed to notify us before starting a new one, were available for the duration of the study, and were able to provide informed consent. We excluded individuals who participated in a chronic pain self-management program within the last year or those who were scheduled for surgical treatment within 6 months. Following the assessment for eligibility, a research team member explained the study procedures and recorded verbal informed consent.

Intervention

The codevelopment of the APM self-management program (thereafter APM or the program) is detailed [32] and available in the study by Marier-Deschenes et al [33]. Briefly, we designed a cognitive behavioral therapy (CBT)–centered, web-based self-management program that would enable participants to develop their self-management skills autonomously (eg, goal setting), practice suggested self-management strategies (eg, pacing), and sustain new behaviors (eg, respect of limits). Despite the diverse nature of our targeted population, the proposed self-management strategies to explore and develop are mostly universal, spanning areas such as managing thoughts and emotions, gradually resuming physical activity, practicing pacing, and adopting good sleep hygiene. The program is self-guided (ie, unguided) in that it provides the same information as face-to-face programs offered in tertiary pain clinics but without therapeutic support from health care professionals [34]. It is structured around weekly lessons over an 8-week period with content that encompasses 26 different behavior change techniques [35] and a downloadable personal plan (Table 1).

Table 1. Summary of Agir pour moi (APM) program's topics and self-management strategies with associated content.

Week	Topics and strategies	Lesson headers
Foreword	What does APM offer?	• What is self-management? • Who is this program for? • Will you have less pain? • Key attitudes to adopt • How to navigate the program?
Week 1	Introduction	• What is chronic pain? • Are you ready for self-management? • How to set specific, measurable, appealing, realistic, and time-bound (SMART) objectives?
Week 2	Engage in well-being activities	• Reduce stress • Experience mindfulness
Week 3	Practice pacing	• Follow-up on last week's objective • Evaluate your energy expenditures
Week 4	Practice pacing, continued	• Follow-up on last week's objective • Planning your weeks
Week 5	Engage in physical activity	• Follow-up on last week's objective • Stretching exercises • Engage in physical activity that you enjoy
Week 6	Take care of your thoughts	• Follow-up on last week's objective • Identify thinking traps • Perceive the positive
Week 7	Revise your lifestyle habits	• Follow-up on last week's objective • Promote sleep • Adapt your diet
Week 8	Plan for the future	• Reflect on previous objectives and further goals • Sustain the change

The program incorporates a variety of media, including photos, infographics, interactive scenarios, tables, audio recordings, and videos. All content is fully narrated, with short audio clips accompanying written information in each lesson, ensuring accessibility in both formats. Interactive exercises such as quizzes, drag-and-drop questions, and real-life scenarios enhance understanding.

Participants were encouraged to set and track their own weekly objectives related to the topic in their personal plan, which serves as our learner's workbook. This plan features reflective, observational, monitoring, problem-solving, and action-planning activities.

Participants were advised to allocate 60 to 90 minutes weekly for program activities. They had the flexibility to divide the lessons into multiple short sessions; completing each lesson in 1 sitting was not necessary. While participants were encouraged to follow the program sequentially, all 8-week content was readily accessible.

XSL•FO
RenderX

Following the program poses minimal health risks. The program incorporates low-intensity activities, such as stretching exercises, which might cause temporary discomfort when resumed. However, the risks associated with physical inactivity, including the development or worsening of chronic illnesses, outweigh those of gradually resuming physical activity.

Quantitative Data Collection and Outcomes

Overview

Participants were assigned a log-in user ID and password for the program's web-based platform. They completed self-reported questionnaires on the web at 3 time points: preintervention, postintervention, and 3 months after completing the program. We sent an email reminder to those who did not log in at least once a week or complete questionnaires at the appropriate time. Participants who completed all questionnaires were eligible for a random computerized drawing of 5 CAD $75 (US $55.89) gift cards. We provided participants facing technical difficulties with phone support. We used REDCap (Research Electronic Data Capture; Vanderbilt University), a secure web application, for creating and managing surveys and databases.

As this was a feasibility study, we did not perform a power calculation on measures of effect but rather aimed at estimating the number of eligible participants and the potential recruitment rate from the center of expertise in chronic pain management waitlist. Therefore, this study is not appropriately powered to assess APM's efficacy [36].

Feasibility and Acceptability Outcomes

We considered the following outcomes in assessing the feasibility of the intervention and research methods and the acceptability of the program: (1) feasibility of recruitment (number of referred adults who responded to the invitation and consented to participate in the study and number of interested adults excluded based on inclusion and exclusion criteria), (2) feasibility of data collection (rate of response to and completion of the questionnaires at each time point), (3) acceptability for those who engaged with the program (mean score to the Acceptability eScale, which includes dimensions of usability and satisfaction) [37,38], and (4) engagement (number of lessons completed). Participants completing at least 6 (75%) out of the 8 weekly lessons were defined as program completers.

Effects Measures

We opted for the French versions of the following self-reported measures based on the recommendations of the Initiative on Methods, Measurement, and Pain Assessment in Clinical Trials [39] (Table 2).

Table 2. Self-reported measures.

Measures	Items, n	Constructs	Score range	High score meaning
Pain Self-Efficacy Questionnaire [40,41]	10	Individual's confidence to perform activities while experiencing pain	0-60	Better self-efficacy
Pain intensity subscale of the Brief Pain Inventory [42,43]	4	Worst, least, average, and current pain intensity	0-10	Worse pain intensity
Pain interference subscale of the Brief Pain Inventory	7	Impact of pain on general activity, mood, walking ability, normal work, sleep, relationships, and enjoyment of life	0-10	Worse pain interference
Anxiety subscale of the Hospital Anxiety and Depression Scale [44]	7	State of anxiety	0-21	Worst anxiety symptoms
Depression subscale of the Hospital Anxiety and Depression Scale	7	State of depression	0-21	Worst depressive symptoms
Pain Catastrophizing Scale [45]	13	Catastrophic thinking and maladaptive responses to pain	0-52	Worst catastrophizing
Patient Global Impression of Change Scale	1	Patient's rating of overall improvement	1-7	Greater change

Statistical Analyses

We performed descriptive statistics using means (SD) for continuous outcomes and frequencies (%) for categorical outcomes. We compared pre-, post-, and follow-up intervention scores for effect measures using repeated-measures linear models. All statistical analyses were conducted using R software (version 4.3.0; R Foundation for Statistical Computing) in RStudio (version 2023.06.0; Posit PBC) [46,47]. Models were fit using lme4 (version 1.1-33; R Foundation for Statistical Computing) [48]. Model fit evaluations and assumption checks were done through visualizations using performance (version 0.10.4) [49]. Effects were considered significant when the 95% CI for the estimates did not include 0.

Qualitative Data Collection and Outcomes

Overview

We formed a heterogeneous group of 20 potential participants using their Acceptability eScale and pre- and posteffect measure scores. We invited participants with positive and negative impressions of the intervention and those for whom we could observe the effects on functioning or not. From a practical perspective, we made the decision to not interview individuals who did not engage with the program at all, as they would not have been able to provide valuable insights into the program's

XSL•FO
RenderX

acceptability and feasibility. However, we conducted interviews with participants who did not complete the program; they were just not specifically selected based on this criterion. We conducted semistructured, audio-recorded, 40-minute phone interviews 5 to 7 months after the intervention. We achieved data saturation with 15 interviews (12/15, 80% women and 3/15, 20% men) and did not deem it necessary to conduct further interviews [50].

Data Analysis

We analyzed the transcriptions using inductive and deductive thematic analysis based on the motivational model for pain self-management [51]. The lead author read the interviews multiple times to obtain a detailed understanding, then coded them according to the research questions and with consideration for the model's components. According to the model by Jensen et al [51], the willingness to embrace pain self-management behaviors is influenced by 2 primary factors. First, it is molded by beliefs concerning the perceived importance of these behaviors, encompassing considerations of cost/benefit ratio, learning history, and current contingencies. Second, self-efficacy, denoting personal beliefs about one's abilities to accomplish a specific task, also plays a pivotal role in shaping the inclination toward behavior change. Furthermore, to ensure the validity of the analysis, a research associate coded and discussed 3 interviews. Then, the lead author further identified meaningful units and assembled them into descriptive categories. She analyzed, interpreted, and summarized categories into 3 explanatory themes that were then discussed among all coauthors [52].

Integration

We addressed our study's main objectives by integrating the quantitative and qualitative results, drawing on all relevant data. We opted for 2 approaches: a weaving approach through the narrative and 2 joint displays presenting categories and associated quotes explaining the quantitative data [53,54].

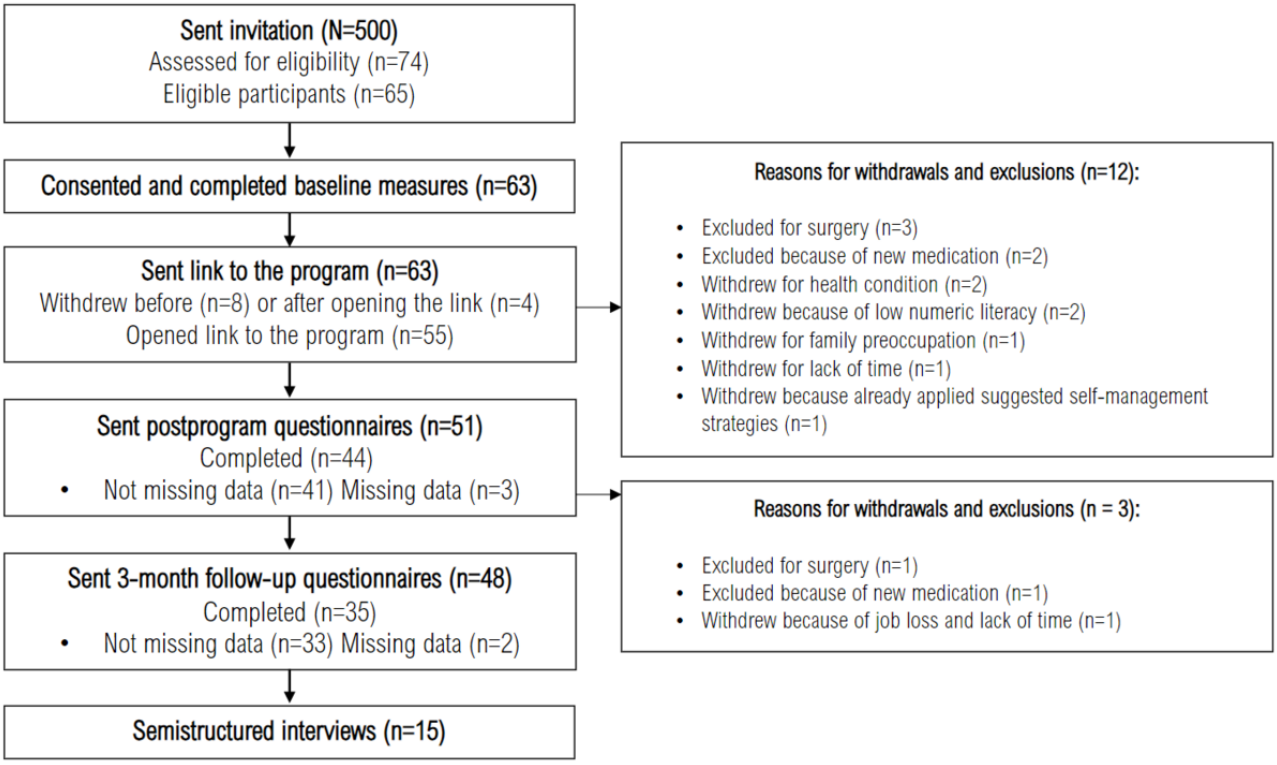
Results

Feasibility

From the 500 invitations sent, 74 (15%) individuals expressed interest, of which 65 (13%) were confirmed eligible. Of these 65 eligible individuals, 63 (97%) consented to participate (Figure 1). A total of 9 (12%) of the 74 participants were ineligible due to unavailability during summer, current services from a pain clinic or rehabilitation center, lack of belief in program helpfulness, absence of computer access, low literacy level, and being unreachable. Response rates were 70% (44/63) at postintervention and 56% (35/63) at 3-month follow-up. A total of 15 (24%) of the 63 participants exited the study for reasons mostly unrelated to the intervention. Missing data were attributed to connectivity problems with the REDCap platform.

Participants were almost exclusively White (61/63, 97%) and female (44/63, 70%), with a mean age of 54 (range 24-75) years. On average, participants experienced chronic pain symptoms for 12 (SD 13.6) years, and most (34/63, 54%) had chronic musculoskeletal pain (Table 3).

Figure 1. Flow diagram of the participants' recruitment, enrollment, and engagement.



XSL•FO
RenderX

Table 3. Participants' sociodemographic and clinical characteristics (N=63).

Sociodemographic characteristics	Values, n (%)
Age (years)	
20 to 29	2 (3)
30 to 39	6 (10)
40 to 49	13 (21)
50 to 59	17 (27)
60 to 69	19 (30)
70 to 79	6 (10)
Gender	
Women	44 (70)
Men	17 (27)
Prefer not to answer	2 (3)
Race	
White	61 (97)
People of color	2 (3)
Marital status	
Married	26 (41)
Living common law	20 (32)
Single	10 (16)
Widowed	3 (5)
Divorced or separated	4 (6)
Education level	
No certificate, diploma, or degree	3 (5)
High-school diploma or equivalency certificate	17 (27)
Apprenticeship, trades certificate, or diploma	6 (10)
College, CEGEP, or other nonuniversity certificate or diploma	20 (32)
University certificate or diploma below bachelor level	5 (8)
University diploma or degree at bachelor level or above	12 (19)
Employment	
Full-time	18 (29)
Part-time	6 (10)
Off work for a short or undetermined term	8 (13)
Off work for a long term or disabled	12 (19)
Unemployed	2 (3)
Retired	16 (25)
Other	1 (2)
Household income	
CAD 0-$49,999 (US $0-$37,257)	19 (30)
CAD $50,000-$99,999 (US $37,258-$74,515)	24 (38)
CAD ≥$100,000 (US ≥$74,516)	12 (19)
Do not know	2 (3)
Prefer not to answer	6 (10)
Duration of chronic pain (years)	

XSL•FO
RenderX

Sociodemographic characteristics	Values, n (%)
1 to 5	26 (41)
6 to 10	15 (24)
11 to 15	6 (10)
16 to 20	5 (8)
≥21	11 (17)
Type of chronic pain	
Chronic widespread pain (includes fibromyalgia syndrome)	24 (38)
Complex regional pain syndrome	8 (13)
Chronic musculoskeletal pain (eg, cervical pain and low back pain)	34 (54)
Chronic headache and orofacial pain (eg, migraine)	10 (16)
Chronic visceral pain	5 (8)
Chronic neuropathic pain	9 (14)
Inflammatory arthritis (eg, rheumatoid arthritis and psoriatic arthritis)	8 (13)
Osteoarthritis	22 (35)
Other	6 (10)

[a]CEGEP : College of General and Professional Teaching.

Acceptability

Among the 63 participants, 38 (60%) completed the Acceptability eScale. The average total score for these participants was 25.2 out of 30 (acceptability threshold=24). In Table 4, we have presented the integrated results of the quantitative and qualitative phases. The mean score for each of the 6 items from the Acceptability eScale is listed in the first column. Each item score ranges from 1 to 5, with higher scores reflecting higher acceptability. The interviews revealed a consistent theme: a program that is globally acceptable, easy to use, and engaging. Interviewees felt they could opt for what worked for them once they had experienced it completely. They would keep most of the program as is, with some potential minor improvements.

XSL•FO
RenderX

Table 4. Joint display of participants' perceptions of the Agir pour moi program's acceptability.

Quantitative results	Qualitative findings	Integrated analysis

Mean score of the ease-of-use item: 4.0/5 (SD 0.9); 5 (8%) of the 63 participants needed support at least once to log in	• Ease of navigation
	  • "Everything is very easy. We can go forward, we can go back, we can resume, we can close, come back. No, everything is perfect." [INT[a] 7]
	• User-friendly, except for the platform hosting the web-based program
	  • "As I was saying, when you're a participant, and you're logged in, I think it is super user-friendly ... It's more when you get to the big link page, I believe that's a little less user-friendly." [INT 13]
		The program was user-friendly once logged in, but connecting to the hosting platform could have been more intuitive.

Mean score of the strategies' helpfulness item: 4.1/5 (SD 0.9)	• Picking the tool you need
	  • "That's it, because if you don't have that resource, what do you do? On a day when things aren't going well, you brood, you grumble all day long; you're in pain, you're angry, and you lose your patience, so if you think a little bit after having followed the program, you go back and find the tool you needed in it. Because as far as I'm concerned, the program wasn't one singular tool; it was a toolkit, and then you take what you need." [INT 10]
	  • "I was happy to see all this in the program, because it helped me a lot personally." [INT 8]
	• Adaptable to one's situation
	  • "It's good to have all the suggestions in there because, ultimately, we keep what suits us and works well for us. So there's something for everyone, for everyone's tastes, and it's not the same stuff that works for everyone." [INT 12]
	• Developing further the strategies that suit you best
	  • "It's been like a springboard to the rest of my journey, to do some reading... ... it sure helps with the perception of managing what we're capable to manage on our own." [INT 6]
		Some strategies benefited many participants, while others only reached a minority. Interviewees latched on to at least 1 lesson that triggered something in them. They acquired something useful out of it, although not all appreciated the same strategy or strategies.

Mean score of the required time item: 4.2/5 (SD 0.7)	• Objective-dependent efforts
	  • "I didn't think it was very demanding, and after that, the rest is up to you... then you do it at your own pace, so I thought it was very appropriate." [INT 13]
	• Motivating and minimal effort required
	  • "Oh, for me, it wasn't that much effort, no. Perhaps the first week's lessons were a little longer than the others that followed. But no, it's really not...it didn't take me any effort, no. It was motivating." [INT 7]
	• A little demanding yet tangible
	  • "I think that's what makes it interesting. Because just reading is okay, but afterward, when you do the exercise, you push yourself a bit on the spot; sometimes, you don't really feel like doing it, but it makes it more tangible." [INT 11]
		The time and effort required to follow the program and apply self-management strategies were adequate, but interviewees highlighted the necessity to reach a certain degree of readiness to change because taking action required some investment.

XSL•FO
RenderX

Quantitative results	Qualitative findings	Integrated analysis

Mean score of the use appreciation item: 4.3/5 (SD 0.7)	• Adapting use to your schedule
	  • "As I said, we could follow it whenever we wanted. I chose the time of the day when I was in better shape. That was something I thought was very nice, you know, not having to connect at a fixed time." [INT 15]
	• Staying engaged
	  • "It went really well. Then, you know, you'd scroll, then you'd click, scroll, click, so that, you know, it kept you present; you had to be there... It's not like a video you start and then lose focus. That's very interesting too." [INT 11]
	• Fun program
	  • "it's super fun." [INT 13]
		Participants appreciated navigating the program at their convenient time. It was fun to use, and the need to scroll through the content kept them engaged.

Mean score of the comprehensibility item: 4.5/5 (SD 0.6)	• Very simple explanation • "It's very...it's simple. It's very well explained." [INT 7]	The program presented well-explained, easy to understand information.

Mean score of the satisfaction item: 4.1/5 (SD 0.7)	• Participants would recommend the program
	• "I really enjoyed the program. I liked that there were videos, it made it more dynamic, I thought they were well done, well constructed ... I'd definitely recommend it." [INT 12]
	• Down-to-earth expectations, no promises
	  • "I thought the program was interesting... Knowing what's in it, I'd do it again today... In the beginning, you don't make any promises. In the program, it says 'learning to live with chronic pain', but there's no promise; it doesn't say: 'Hey, when you get to the eighth week, you're pain-free'." [INT 10]
	  • "But I think the program doesn't apply to me... You know, it's like, geez, I thought I would discover something amazing. It's been four years now, I've seen three internists, lots of doctors, and I went to a maxillofacial specialist back in September, and we're trying to figure it out, but nobody knows what it is. I'm still waiting to find out." [INT 2]
		All interviewees mentioned they would recommend the program to someone in a similar situation, reflecting satisfaction. Participants with high expectations for very specific problems might have been less satisfied.

Total mean score of the Acceptability eScale[b]: 25.2/30 (SD 3.0)	• No change required
	  • "Well, it went well. I liked it a lot, I really liked the way it was put together, the way it was presented, gradually if you like. The program is super well done, I'm going to revise it, but I wouldn't change a thing if you asked me if there was anything to change." [INT 1]
	• Use of multiple media
	  • "I loved it because there were testimonials. It wasn't just reading. I put on my headphones, and I didn't have to read. I listened. I like to listen, and then as I filled out my sheet, I'd make notes on the paper as I listened. Sometimes I'd take breaks, put it on pause and come back." [INT 9]
	• Promoting access for everyone
	  • "God, yes, it's acceptable, and everyone should have access to it." [INT 7]
		Overall, the program was well developed for our target users and proposed an appreciated gradual approach to the application of different strategies.

Of the 15 interviewees, 10 (67%) were still accessing the program or using the personal plan on an occasional to regular basis for >5 months after they finished it for the first time.		While the web-based format did not suit everyone, participants appreciated using it whenever they wanted and having the possibility to go back to previous lessons for a refresher. Not traveling to learn self-management strategies was a significant advantage. Had the program been offered through weekly classes at a specific time and place, some participants would have been unable to drive and attend.

XSL•FO
RenderX

Quantitative results	Qualitative findings	Integrated analysis

	• Web-based content accessibility
	  • "It helps a lot that it's online, so you can do it when it suits you. Because for me, if I'd had to travel, you know, appointments and things like that, I'd have had a lot more trouble because it's hard for me to go on the road, whereas here, it was much more accessible, which is really great." [INT 12]
	• Possible punctual use once the program is done
	  • "Then again, it's not a program that once it's done, it's done, and you can't go back to it; you can go back and look. So you can continue to use it." [INT 10]	

[a]INT: interviewee.

[b]The Acceptability eScale has 6 items with a total score ranging from 6 to 30. Higher scores represent a high level of acceptability.

Engagement

Of the 63 participants who consented to the study, 46 (73%) started the program, 29 (46%) finished at least 6 weeks' lessons, and 26 (41%) completed all the lessons (Table 5).

Participants generally followed the lessons in order. Among the 19 noncompleters who did not withdraw from the study, 10 (53%) wrote back to us after receiving the email reminder. These participants provided personal reasons for not following the program according to schedule, including sickness or mental health issues among close relatives, the death of a parent, estate management, and increased symptoms. Including participants who withdrew from the study, 21 (33%) out of 63 participants were not connecting weekly or stopped at some point for reasons external to the study or the intervention. Reasons related to the program included having already applied the proposed strategies and having trouble connecting. Of the 63 participants, 5 (8%) needed help logging in independently (n=3, 60% only required information by email and n=2, 40% received phone support and succeeded with a step-by-step explanation but did not connect afterward).

Table 5. Participants' weekly lessons' level of completion (N=63).

Weekly lessons	Completed the lesson, n (%)	Partially completed the lesson, n (%)	Never initiated the lesson, n (%)
Lesson 1	42 (67)	4 (6)	17 (27)
Lesson 2	37 (59)	6 (10)	20 (32)
Lesson 3	33 (52)	6 (10)	24 (38)
Lesson 4	32 (51)	3 (5)	28 (44)
Lesson 5	30 (48)	5 (8)	28 (44)
Lesson 6	29 (46)	3 (5)	31 (49)
Lesson 7	28 (44)	4 (6)	31 (49)
Lesson 8	26 (41)	4 (6)	33 (52)

Interviewees mentioned other factors potentially undermining the engagement in completing the program. Interviewees highlighted that timing and pain acceptance played an essential role in their perseverance or lack thereof. For example, understanding that no specific cause for her pain issue might ever be found was a turning point in a participant's proactivity:

> *When I was still at the stage where I thought we would find a cause, it is something I really wouldn't have been ready to do, the program... Doing the program, really, it sort of happened right at the time I got to the stage of realizing: 'Okay, now I'm going to stay that way. What am I going to do with this?' I felt a lot of psychological distress. Then, looking more into this (psychological) aspect through the program, it was as if I needed to do this. I'm not done, but it's come a long way. And also, time matters; getting used to accepting.* [INT 11]

For some interviewees, occasional or regular coaching would have been necessary to sustain motivation, answer questions, and revise their personal plan. Of the 63 participants, 2 (3%) mentioned their need for support and feedback:

> *There were ups and downs. There was no one to answer my questions. I found that very, very hard, especially the first few weeks.* [INT 5]
>
> *I really need someone who'll say, "Go, we'll do this, we'll do that." In writing, I'll read, and I'll say to myself, "Oh my god, that's wonderful," but I'll do it 2 or 3 times, and then I'll give up... But I've always been like that; I've always needed someone to push me, not in everything, but especially since I've been ill.* [INT 9]

There are participants who related less to parts of the program because their activity level was not adequately represented:

XSL•FO
RenderX

> *We always assume that people aren't physically active when in pain. Or if they are, clearly, it can't be too much. But you know, on my part, I will get injured before I stop [laughs].* [INT 11]

Furthermore, interviewees described factors mostly contributing to their engagement in completing the program. They expressed a connection to the program's content and felt more hopeful, supported, and less alone listening to the multiple integrated testimonials of individuals with lived experiences applying the strategies:

> *For me, it was seeing lots of testimonials from many people ... that's what really attracted me. To understand it better and to see that that's how it is: everyone's gone down exactly the same path I went through, and we all arrive at the same point, not being able to get out, not seeing anyone. It seemed like it was just me who was going through this. So, it gave me a boost. It also gave me a bit of confidence... I had the impression of being accompanied.* [INT 14]

Moreover, a participant living far from a major urban area highlighted the appeal of these short videos:

> *Here, in my neighborhood, there's no program like that. I know there are places, there are meetings for people with chronic pain to chat, have a coffee, and things like that. We don't have that here. I found it fun to listen to them and know we're not alone in this.* [INT 9]

On the one hand, interviewees thought the weekly connection and review of personal objectives supported their motivation and helped them stay focused:

> *I guess I needed to be held by the hand for a while, and to be guided through it... There was some kind of follow-up. So you weren't left to your own devices as much.* [INT 12]

On the other hand, it might have been too much for some individuals:

> *There were too many things. Every week there was something to do, you know. You didn't have time to swallow the information and were already moving on to other things... Maybe it would have taken a week or two between each chapter. What made it easier was to stop for a few weeks, then think about all the information.* [INT 5]

There is a fine line between what feels like a comprehensive program and what feels like an overwhelming task for unsupported participants. While most interviewees showed interest in all the presented topics and perceived the lessons' sequence as logical, gradual, and positive, specific strategies within these lessons might have lacked appeal to some participants. A suggestion was made that participants could start with their most appealing lessons after covering the pain education section. While we invited everyone to complete the tasks in order, the content was freely available and one could decide to skip parts of the program if desired.

Through the interviewee discourse, we could distinguish microlevel engagement, including the number of lessons they have followed, from macrolevel engagement, referring to the depth of involvement with the behavior change process, such as applying strategies consistently in a real-world setting [55,56]. Behavior change could be challenging, with participants facing expectations from others and their own. However, the program provided them with a better understanding of why it is beneficial to do so, and some participants made it a priority. Getting into the habit of doing something sometimes required broader or prior changes such as setting boundaries with the extended family. The notion of beginning with the easiest tasks was mentioned to underscore the initial challenge of implementing a strategy. Interviewees described the integration of certain habits into their routine as gradual, sometimes evolving over several months, without the participants realizing it. Over 5 months after finishing the program, interviewees described what they continued to apply and how some strategies became part of their routine:

> *Over the weeks, I really selected what had a positive impact on me. Then I do it regularly, almost every day. I'm into meditation, cardiac coherence, managing energy, stretching, and the gratitude journal. The other things, I can't think of anything else I could do more.* [INT 6]

> *I still do so to this day, which is unusual for me. So, I thought it was really, really good... I'm more active now too, I do my exercises almost daily. I used to say to myself, "It's no use, I won't be able to do it. I don't have the energy. I'm in pain. I'm out of shape." Now I say, "Look, do it, even if it's just two minutes today, it'll at least be two minutes, you'll have done it." [INT 12]*

Preliminary Effects Outcomes

We observed an indication of improvements over time in self-efficacy, pain interference, depression, and pain catastrophizing between baseline and postintervention (Table 6). The results of the linear model corrected for the individual (treatment effects) are presented in Table 7. In addition, 24 (56%) out of 43 participants reported some level of improvement on the Patient Global Impression of Change Scale after the intervention.

From a qualitative perspective, 14 (93%) out of 15 interviewees reported a different set of effects 5 to 7 months later. These effects included, among others, shorter and less frequent pain attacks, better management of pain, higher sense of control, less comparison with life before pain, improved psychological state, more patience, less frustration and irritability, forgiveness toward oneself, higher activity level, more self-care, and increased social life.

In Table 7, we have presented the integrated results of the quantitative and qualitative phases, with the overarching theme that the various effects observed were potentially influenced by macrolevel engagements.

The effects of the lifelong task of self-management might become noticeable over time. A participant mentioned

XSL•FO
RenderX

recognizing these effects several months after completing the program:

> *While I was going through (the program), well, you know, it was okay. In my case, it's really like it slowly permeated me, but in a positive way, I mean.* [INT 11]

Therefore, conducting the interviews 5 to 7 months after program completion was deemed appropriate.

Table 6. Mean scores for outcome measures at baseline, postintervention, and 3-month follow-up.

Measure	Subscale or construct	Baseline (N=63), mean (SD)	Postintervention (n=43), mean (SD)	3-month follow-up (n=34), mean (SD)
Pain Self-Efficacy Questionnaire[a]	Self-efficacy	26.9 (13.8)	32.4 (13.3)	32.5 (12.4)
Brief Pain Inventory	Pain intensity	5.5 (1.6)	4.8 (1.5)	5.0 (1.6)
Brief Pain Inventory	Interference	5.7 (2.0)	4.4 (2.0)	4.7 (1.9)
Hospital Anxiety and Depression Scale	Anxiety	9.0 (4.4)	8.4 (4.2)	8.2 (4.0)
Hospital Anxiety and Depression Scale	Depression	8.5 (4.0)	7.2 (3.9)	7.5 (3.8)
Pain Catastrophizing Scale	Catastrophizing	25.3 (10.9)	20.1 (12.2)	20.5 (13.2)

[a]A higher score in the Pain Self-Efficacy Questionnaire indicates a high level of self-efficacy.

XSL•FO
RenderX

Table 7. Joint display of participants' perceptions of the Agir pour moi (APM) program's effects.

Quantitative results	Qualitative findings	Integrated analysis

• Pain Self-Efficacy Questionnaire: the scores increased by 3.06 points on average after intervention (95% CI 1.26 to 4.85), *P*=.002.
• No effects were detected when examining postintervention and follow-up metrics.
	• Better self-efficacy
	  • "I'm less inclined to compare myself with my previous life. I see more of what I'm capable of doing today with the abilities I have. That's the main thing, I think." [INT[a] 12]
	  • "Even though I've seen specialists who gave me medication, in the end, I think I got better by doing this on my own and saying to myself, Okay, I'm basically taking charge... I feel like I can control my pain peaks a bit more, and I know why I will have them." [INT 4]
	  • "Let's just say I've relearned how to gain confidence in myself and then say you're capable, go ahead, go take a walk, you can do it." [INT 8]
		Participants' belief in their capacity to do certain things to achieve their goal increased.

• BPI[b] pain intensity subscale scores decreased by 0.32 points on average after intervention (95% CI –0.55 to –0.10), *P*=.007.
• No effects were detected when examining postintervention and follow-up metrics.
	• No trend in pain intensity changes
	  • "I can say that my pain is less present and that it's not all I think about anymore. So yes, for my pain, it helped a lot." [INT 5]
	  • "It's just that instead of being in pain for nine hours at a time, well, not only have I hardly had any attacks for six, seven months, but when it happens, well the two times it happened, it lasted two, three hours or so, then it's stopped really suddenly, instead of lasting a whole night. So I tell you, it's not so bad." [INT 14]
	  • "But the pain remains the same. Sometimes when I'm doing your stuff, I've managed to get away for half an hour or an hour, but it comes right back." [INT 3]
	  • "I've done things, like mediation and all that, but it doesn't work." [INT 2]
		There were improvements for some interviewees and no improvements for others. While not measured with the BPI subscale, the frequency of pain crises decreased in some participants. However, as mentioned straightforwardly at the beginning of the APM, pain reduction is not the main objective of a self-management program.

• BPI pain interference subscale: scores decreased by 0.78 points on average after intervention (95% CI –1.19 to –0.37), *P*=.001.
• No effects were detected when examining postintervention and follow-up metrics.
	• Less interference
	  • "I realize that the energy is coming back a little. It's coming back. It's not a complete crash like it used to be for me, and then I'd take a week to recover because I'd gone over the edge. Now, I don't do that anymore ... And in the end, I do a lot more than I used to do that way." [INT 12]
	  • "You know, I was more or less able to do some movements, some I couldn't do at all anymore. There are some that I've gradually managed to recover a little, it's not to the maximum here, but ... like putting on my shoes." [INT 15]
	  • "Well, I fall asleep faster when I do these exercises. Because I always try to go to bed at the same time, and sometimes sleep doesn't come. So I tell myself I'm going to bed anyway and do some breathing exercises. Then I fall asleep, which doesn't take long." [INT 15]
		Reducing pain interference translates into a gain in energy, a more stable ability to perform tasks, a new capacity to do movements, increased social activities, and better sleep.

XSL•FO
RenderX

Quantitative results	Qualitative findings	Integrated analysis

• HADS[c] depression subscale: scores decreased by 0.73 points after intervention (95% CI –1.21 to –0.25), *P*=.005.
• No effects were detected when examining postintervention and follow-up metrics.
	• Positive change in depressive state
	  • "It had been seven years since I'd stopped putting any effort into it and let myself fall, so it...no, no, it whipped me, and then I seemed to become a bit like myself again. I was letting myself go, then it was like: okay, go, I've been sinking for seven years, and now it's time to get back on. It gave me a good boost. ... My mood has changed. I seem to be less, sorry about my French, but I'm less (swear) angry all day long." [INT 14]
	• Better mood
	  • "how can I put it, patience, my patience came back, better than when I couldn't do my things. Yes, yes, that, I've made some gains." [INT 15]
	  • "It's also psychological, you know, like being less on edge in my head, I was a lot like (swears), I can't do this anymore, I can't. So I'm in a better mood with the kids. ... It also has a lot to do with irritability, because you know when you're in pain, you're always irritable, so if I'm in less pain, I'm less irritable, and I'm more likely to want to go outside with them." [INT 11]
		Most interviewees did not explicitly talk about depression but many reported being in a better mood, being less frustrated, and being less irritable.

• HADS anxiety subscale: nonsignificant score decrease of 0.37 points after intervention (95% CI –0.85 to 0.12), *P*=.14. • No effects were detected when examining postintervention and follow-up metrics.	• Anxiety • "it's really taken my anxiety level about it down a notch." [INT 11]	Following the program does not appear to impact the anxiety state. Only 1 interviewee briefly mentioned a decrease.
• Pain Catastrophizing Scale: scores decreased by 2.83 points after intervention (95% CI –4.35 to –1.30), *P*=.001. • No effects were detected when examining postintervention and follow-up metrics.	• Less panic • "For me, what was a real game changer ... was also realizing that most people in this condition tend to have the same thoughts; thinking, for example, that there's something perhaps serious hidden behind it, thinking that it will never stop, that you could die from it, and all that. It was good for me, because these are patterns I really have. Then, I thought of myself more like a normal person. ... The panic I used to feel about my pain has almost completely disappeared, and I realize how much I can change my situation myself." [INT 11]	The testimonials and theoretical content helped normalize some participants' catastrophic thoughts and guide them in confronting those and adopting more adapted views. Following the program can reduce catastrophizing.

[a]INT: interviewee.

[b]BPI: Brief Pain Inventory.

[c]HADS: Hospital Anxiety and Depression Scale.

Discussion

Principal Findings

This paper described a pilot, mixed methods study assessing feasibility and acceptability and exploring the preliminary outcomes of a new self-directed, web-based program for chronic pain among adults awaiting superspecialized services. Collaborating with patient partners experiencing diverse types of pain was a key strategy enabling us to align the program closely with the varied needs and expectations of our wide-ranging audience. We opted for a simple yet attractive layout, providing clear instructions and features to aid those with attention and concentration challenges.

Feasibility, Acceptability, and Engagement

Chronic noncancer pain affects individuals in different ways and to different degrees. Those awaiting tertiary services in Canada experience severe impairments and present with a poor biopsychosocial profile [2]. To specifically recruit these individuals who are not yet patients at the center of expertise in chronic pain management, we could not directly reach out to them. The hospital's archives had to send them invitation letters. Estimating the response rate in such a specialized context proved challenging because we lacked a benchmark for our expectations. However, a study by Thiblin et al [57], which involved a comparable internet-administered, CBT-based self-help intervention, achieved an 11% enrollment rate by

XSL•FO
RenderX

sending out 509 invitation letters. Considering the documented high dropout rates in similar trials [19,58], we anticipated that 500 invitations would suffice to ensure a minimum of 30 participants responding to questionnaires at all 3 time points. Our enrollment rate for a 3-month recruitment period was similar to or better than those of studies recruiting in tertiary pain treatment facilities [59,60]. The consent rate among potentially eligible adults was acceptable. However, because 4 (6%) out of 63 participants exited the study for surgery purposes and 3 (5%) others for significant changes in their medication, we might need to reconsider how we address these eligibility criteria during the phone interview. Furthermore, considering there are between 1000 and 1500 individuals awaiting services from the center of expertise and that they all would not meet our eligibility criteria, our response rate, although similar to those observed in other studies, would require conducting a future full trial across multiple centers.

Our study yielded comparable data collection results to other web-based intervention pilot trials with approximately 50% (35/63) of response at 3-month follow-up [61,62]. While studies with higher financial incentives (US $25-$80 per assessment) during visit assessments or initial motivational interviews had better response rates [60,63-66], we purposefully chose to stay as close as possible to real life where no incentives are offered. We received no negative feedback on the number or length of questionnaires. However, sending email reminders to participants who did not log in or complete the questionnaires at the appropriate time was suboptimal. Making phone calls in addition to email reminders might have provided us with reasons for disengagement and ensured participants received the reminders.

Some interviewees highlighted that the log-in process was not intuitive, leading us to consider modifying this aspect before conducting a full trial. The log-in was essential to the research project but is not part of APM itself and will not apply in real life. Once logged in, accessing APM at their most convenient time and place was a significant asset for our participants, consistent with patients' preferences [21].

APM's codevelopment with health care professionals and people with lived experience of chronic pain allowed a tailored approach to this population's needs and preferences in a web-based self-management program [32]. The interviewees' description of APM aligned with the acceptability score as a globally acceptable, easy-to-use, and engaging program. We based our 24 (80%) out of 30 threshold score for the Acceptability eScale based on the study by Tariman et al [38], suggesting that 80% of the highest possible summary score indicates good program acceptability. However, a score <24 would not automatically deem the program unacceptable. We must examine individual item scores to assess specific program weaknesses. All items scored ≥4 (≥80%) out of 5, as show in Table 4, indicating no significant flaws requiring major modifications. Minor improvements we could make include adding testimonials from highly active individuals with chronic pain who learned to pace themselves and mentioning that lessons are preferably followed in order but can also be explored based on personal preferences after completing week 1. APM effectively promoted behavioral change, guided participants in taking action, and served as a reference in the longer term.

While participants' weekly lessons' level of completion was lower than anticipated, it was consistent with what had been observed in other feasibility and pilot studies [62,67-69] and could be explained mainly by reasons external to the study. Our qualitative results alleviated concerns about potential flaws and did not point toward questioning the participant's appeal to the program. Overall, we are confident that the program and trial procedures are both feasible and acceptable.

Preliminary Effects

We explored pre- and postintervention effects as preliminary indications of potential changes in self-efficacy, pain intensity and interference, anxiety, depression, and catastrophizing. Findings yielded relevant results, but these should be interpreted cautiously.

Nevertheless, the qualitative interviews pointed in the same direction as our preliminary quantitative findings. Furthermore, these aligned with the results of a meta-analysis suggesting that following internet-delivered, CBT-centered interventions for chronic pain can lead to small significant improvements in pain interference and intensity, depression, anxiety, self-efficacy, and catastrophizing, with greater treatment effects in anxiety, pain interference, and intensity in guided compared to unguided interventions [34]. Therefore, depending on their perceived importance of change and self-efficacy, individuals with chronic pain may require additional support in reaching readiness to make sustainable changes.

Because we still did not know precisely what clinical, intervention, and study characteristics positively impacted the effects of unguided CBT-based self-management programs for chronic pain, APM offered several self-management strategies [70]. However, we did not expect participants to implement all of them once they completed the program. Participants used this program as a toolbox, as mentioned by an interviewee.

No adverse events were reported throughout the course of this study.

Limitations

This study has some limitations. First, we cannot make definitive statements regarding APM's effects without an appropriate control group, randomization process, and sample size. Indeed, we neither designed nor appropriately powered this feasibility trial to test a specific hypothesis [36]. Furthermore, while the participants presented various pain conditions, most of the female participants were White, as this is the case in similar trials [18,71,72], and were all attached to a single center. Recruitment strategies of a future randomized controlled trial should focus on attracting a broader representation of individuals with chronic pain in terms of gender, ethnicity, and health care institutions. In Quebec City, where our study was conducted, <10% of the population identified as members of a visible minority group in 2021. Expanding our research to cities with greater ethnic diversity could enhance our sociodemographic data and improve the relevance of our findings. Shifting to web-based recruitment methods might allow us to create tailored

XSL•FO
RenderX

invitation messages for specific demographic groups, using a casual design and images, instead of overwhelming potential participants with excessive written information. In our feasibility study, we had to use standardized letters to recruit participants from the waitlist of the center of expertise in chronic pain management. We acknowledge the need to adopt more flexible parameters for future large-scale studies. Furthermore, we might adjust our eligibility criterion, not limiting participation to those awaiting specialized services. This broader criterion could yield a more diverse sample, aligning with our aim to reach a wider demographic. However, we are mindful of the potential impacts on adherence and user satisfaction this broader criterion might pose. Nevertheless, these adjustments reflect our dedication to conducting a comprehensive and inclusive trial, ultimately contributing to a more nuanced understanding of chronic pain management. This may lead us to unforeseen modifications in our program.

Despite widespread internet access in Canada, disparities in internet speed, affordability, and digital literacy persist. APM, being exclusively web-based, poses a limitation in reaching individuals from remote regions and Indigenous communities as well as those in low-income households, older adults, and individuals with disabilities. These groups are disproportionately affected by the digital divide in Canada, making it challenging for them to access our program.

Future Direction

This study provides an initial understanding of APM's potential benefits for this group of individuals with chronic pain awaiting specialized services. Through the interviews, we acknowledged we had not captured the effects on the temporal aspect of pain, such as shorter and less frequent pain attacks, which were crucial for some participants. Therefore, we could consider adding measures capturing these aspects in a future trial.

Developing a web- and evidence-based, patient-centered, free-of-charge, user-friendly, and French self-management program for chronic noncancer pain represents a potential response to the clearly expressed needs of individuals with this condition. Although the literature increasingly emphasizes the importance of personalization in eHealth, our limited financial resources hindered us from incorporating advanced features. We deliberately chose to focus on fundamental aspects and prioritize what we could offer and support in the long term, establishing the groundwork for a web-based program that could potentially evolve. As a result, the current version did not include personalized features, but it was still perceived as usable and useful. It will be essential to document how the program's implementation makes it possible to respond quickly and more equitably to some of the needs of patients waiting for services or who live far from large centers. APM is currently being used without restrictions in other French-speaking regions and countries. Anyone can use it freely, but a potential hurdle faced when using it abroad pertains to adapting to the accents in testimonial videos and Quebec-specific expressions.

Conclusions

The study findings provided preliminary evidence that the APM program and research methods were both feasible, as suggested by perceived acceptability and engagement. Furthermore, it provided preliminary indications of potential improvements in self-efficacy, pain intensity, interference, depression, and catastrophizing. The study yielded essential results to undertake a future complete trial.

Acknowledgments

The authors would like to thank Sergio Cortez Ghio, a biostatistician who performed the quantitative analysis. The authors would also like to extend their thanks to all those who contributed to the codevelopment of APM. PMD acknowledges support from a Frederick Banting and Charles Best Canada Graduate Scholarship Doctoral program from the Canadian Institutes of Health Research, a Leadership and Sustainable Development scholarship from Laval University, and a scholarship from the Strategic Research Development Fund from VITAM—Center for Sustainable Health Research. AMP holds the Chair in Educational Leadership on Chronic Pain—Medisca, which contributed to the development of the program. Additional funding for this study came from discretionary funds. The funders did not play any role in the design, management, analysis, interpretation, or writing of the manuscript.

Data Availability

The data sets generated during and analyzed during this study are available from the corresponding author on reasonable request.

Authors' Contributions

PMD, AMP, and AL contributed to the study conception and design. PMD and AMP supported the details of recruitment and data collection. PMD, AL, and LJ analyzed the results. PMD and AL drafted the manuscript. All authors commented on and revised its previous versions before reading and approving the manuscript to be published.

Conflicts of Interest

None declared.

Multimedia Appendix 1

XSL•FO
RenderX

CONSORT-eHEALTH checklist (V 1.6.1).
[PDF File (Adobe PDF File), 1181 KB - humanfactors_v11i1e50747_app1.pdf]

References

1. Ndegwa S, Dolcine B, Frey N, Quay T. Access to and availability of non-pharmacological treatments for chronic non-cancer pain in Canada: an environmental scan. Canadian Agency for Drugs and Technologies in Health (CADTH). 2018 Nov. URL: https://www.cadth.ca/sites/default/files/pdf/es0313_access-to-non-pharmacological-treatment-of-pain.pdf [accessed 2024-04-18]
2. Choinière M, Dion D, Peng P, Banner R, Barton PM, Boulanger A, et al. The Canadian STOP-PAIN project - part 1: who are the patients on the waitlists of multidisciplinary pain treatment facilities? Can J Anaesth 2010 Jun;57(6):539-548. [doi: 10.1007/s12630-010-9305-5] [Medline: 20393821]
3. Dueñas M, Ojeda B, Salazar A, Mico JA, Failde I. A review of chronic pain impact on patients, their social environment and the health care system. J Pain Res 2016;9:457-467 [FREE Full text] [doi: 10.2147/JPR.S105892] [Medline: 27418853]
4. Burke AL, Mathias JL, Denson LA. Psychological functioning of people living with chronic pain: a meta-analytic review. Br J Clin Psychol 2015 Sep;54(3):345-360. [doi: 10.1111/bjc.12078] [Medline: 25772553]
5. Moriarty O, McGuire BE, Finn DP. The effect of pain on cognitive function: a review of clinical and preclinical research. Prog Neurobiol 2011 Mar;93(3):385-404. [doi: 10.1016/j.pneurobio.2011.01.002] [Medline: 21216272]
6. Reid KJ, Harker J, Bala MM, Truyers C, Kellen E, Bekkering GE, et al. Epidemiology of chronic non-cancer pain in Europe: narrative review of prevalence, pain treatments and pain impact. Curr Med Res Opin 2011 Feb;27(2):449-462. [doi: 10.1185/03007995.2010.545813] [Medline: 21194394]
7. Ashton-James CE, Anderson SR, Mackey SC, Darnall BD. Beyond pain, distress, and disability: the importance of social outcomes in pain management research and practice. Pain 2022 Mar 01;163(3):e426-e431 [FREE Full text] [doi: 10.1097/j.pain.0000000000002404] [Medline: 34252908]
8. Deslauriers S, Roy JS, Bernatsky S, Blanchard N, Feldman DE, Pinard AM, et al. The burden of waiting to access pain clinic services: perceptions and experiences of patients with rheumatic conditions. BMC Health Serv Res 2021 Feb 18;21(1):160 [FREE Full text] [doi: 10.1186/s12913-021-06114-y] [Medline: 33602224]
9. Blanchard N, Deslauriers S, Gervais-Hupé J, Hudon A, Roy JS, Bernatsky S, et al. "It feels like an endless fight": a qualitative study exploring healthcare utilization of persons with rheumatic conditions waiting for pain clinic admission. BMC Musculoskelet Disord 2022 Sep 22;23(1):878 [FREE Full text] [doi: 10.1186/s12891-022-05808-6] [Medline: 36131335]
10. Burke AL, Mathias JL, Denson LA. Waiting for multidisciplinary chronic pain services: a prospective study over 2.5 years. J Health Psychol 2020 Aug;25(9):1198-1212. [doi: 10.1177/1359105317752828] [Medline: 29322830]
11. Choinière M, Peng P, Gilron I, Buckley N, Williamson O, Janelle-Montcalm A, et al. Accessing care in multidisciplinary pain treatment facilities continues to be a challenge in Canada. Reg Anesth Pain Med 2020 Dec;45(12):943-948. [doi: 10.1136/rapm-2020-101935] [Medline: 33024007]
12. Campbell F, Hudspith M, Choinière M, El-Gabalawy H, Laliberté J, Sangster M. Working together to better understand, prevent, and manage chronic pain : what we heard / a report by the Canadian pain task force. Health Canada. 2020. URL: https://publications.gc.ca/site/eng/9.892793/publication.html [accessed 2024-04-05]
13. Campbell F, Hudspith M, Anderson M, Choinière M, El-Gabalawy H, Laliberté J, et al. Chronic pain in Canada: laying a foundation for action: a report of the Canadian pain task force. Health Canada. 2019. URL: https://www.canada.ca/content/dam/hc-sc/documents/corporate/about-health-canada/public-engagement/external-advisory-bodies/canadian-pain-task-force/report-2019/canadian-pain-task-force-June-2019-report-en.pdf [accessed 2024-04-05]
14. Devan H, Perry MA, van Hattem A, Thurlow G, Shepherd S, Muchemwa C, et al. Do pain management websites foster self-management support for people with persistent pain? A scoping review. Patient Educ Couns 2019 Sep;102(9):1590-1601. [doi: 10.1016/j.pec.2019.04.009] [Medline: 30981410]
15. Buhrman M, Syk M, Burvall O, Hartig T, Gordh T, Andersson G. Individualized guided internet-delivered cognitive-behavior therapy for chronic pain patients with comorbid depression and anxiety: a randomized controlled trial. Clin J Pain 2015 Jun;31(6):504-516. [doi: 10.1097/AJP.0000000000000176] [Medline: 25380222]
16. Irvine AB, Russell H, Manocchia M, Mino DE, Cox Glassen TC, Morgan R, et al. Mobile-Web app to self-manage low back pain: randomized controlled trial. J Med Internet Res 2015 Jan 02;17(1):e1 [FREE Full text] [doi: 10.2196/jmir.3130] [Medline: 25565416]
17. Dear BF, Gandy M, Karin E, Staples LG, Johnston L, Fogliati VJ, et al. The pain course: a randomised controlled trial examining an internet-delivered pain management program when provided with different levels of clinician support. Pain 2015 Oct;156(10):1920-1935 [FREE Full text] [doi: 10.1097/j.pain.0000000000000251] [Medline: 26039902]
18. Bromberg J, Wood ME, Black RA, Surette DA, Zacharoff KL, Chiauzzi EJ. A randomized trial of a web-based intervention to improve migraine self-management and coping. Headache 2012 Feb;52(2):244-261 [FREE Full text] [doi: 10.1111/j.1526-4610.2011.02031.x] [Medline: 22413151]
19. Moman RN, Dvorkin J, Pollard EM, Wanderman R, Murad MH, Warner DO, et al. A systematic review and meta-analysis of unguided electronic and mobile health technologies for chronic pain-is it time to start prescribing electronic health applications? Pain Med 2019 Nov 01;20(11):2238-2255. [doi: 10.1093/pm/pnz164] [Medline: 31386151]

XSL•FO
RenderX

20. Moore SK, Guarino H, Acosta MC, Aronson ID, Marsch LA, Rosenblum A, et al. Patients as collaborators: using focus groups and feedback sessions to develop an interactive, web-based self-management intervention for chronic pain. Pain Med 2013 Nov;14(11):1730-1740 [FREE Full text] [doi: 10.1111/pme.12200] [Medline: 23859438]
21. Ledel Solem IK, Varsi C, Eide H, Kristjansdottir OB, Mirkovic J, Børøsund E, et al. Patients' needs and requirements for eHealth pain management interventions: qualitative study. J Med Internet Res 2019 Apr 01;21(4):e13205 [FREE Full text] [doi: 10.2196/13205] [Medline: 30877780]
22. Banner D, Bains M, Carroll S, Kandola DK, Rolfe DE, Wong C, et al. Patient and public engagement in integrated knowledge translation research: are we there yet? Res Involv Engagem 2019;5:8 [FREE Full text] [doi: 10.1186/s40900-019-0139-1] [Medline: 30805202]
23. Karazivan P, Dumez V, Flora L, Pomey MP, Del Grande C, Ghadiri DP, et al. The patient-as-partner approach in health care: a conceptual framework for a necessary transition. Acad Med 2015 Apr;90(4):437-441. [doi: 10.1097/ACM.0000000000000603] [Medline: 25607943]
24. Slattery P, Saeri AK, Bragge P. Research co-design in health: a rapid overview of reviews. Health Res Policy Syst 2020 Feb 11;18(1):17 [FREE Full text] [doi: 10.1186/s12961-020-0528-9] [Medline: 32046728]
25. Kirwan JR, de Wit M, Frank L, Haywood KL, Salek S, Brace-McDonnell S, et al. Emerging guidelines for patient engagement in research. Value Health 2017 Mar;20(3):481-486 [FREE Full text] [doi: 10.1016/j.jval.2016.10.003] [Medline: 28292494]
26. Eccleston C, Blyth FM, Dear BF, Fisher EA, Keefe FJ, Lynch ME, et al. Managing patients with chronic pain during the COVID-19 outbreak: considerations for the rapid introduction of remotely supported (eHealth) pain management services. Pain 2020 May;161(5):889-893 [FREE Full text] [doi: 10.1097/j.pain.0000000000001885] [Medline: 32251203]
27. Ivankova NV, Creswell JW, Stick SL. Using mixed-methods sequential explanatory design: from theory to practice. Field Methods 2016 Jul 21;18(1):3-20. [doi: 10.1177/1525822X05282260]
28. Creswell JW, Clark VL. Designing and Conducting Mixed Methods Research. Thousand Oaks, CA: SAGE Publications; 2007.
29. Bowen DJ, Kreuter M, Spring B, Cofta-Woerpel L, Linnan L, Weiner D, et al. How we design feasibility studies. Am J Prev Med 2009 May;36(5):452-457 [FREE Full text] [doi: 10.1016/j.amepre.2009.02.002] [Medline: 19362699]
30. Lancaster GA, Thabane L. Guidelines for reporting non-randomised pilot and feasibility studies. Pilot Feasibility Stud 2019;5:114 [FREE Full text] [doi: 10.1186/s40814-019-0499-1] [Medline: 31608150]
31. Eysenbach G, CONSORT-EHEALTH Group. CONSORT-EHEALTH: improving and standardizing evaluation reports of web-based and mobile health interventions. J Med Internet Res 2011 Dec 31;13(4):e126 [FREE Full text] [doi: 10.2196/jmir.1923] [Medline: 22209829]
32. Marier-Deschenes P, Pinard AM, LeBlanc A. Codéveloppement d'un programme d'autogestion de la douleur chronique en ligne: un projet de recherche basé sur la conception et axé sur l'engagement des patients. Can J Pain 2024;8(1):2302487 [FREE Full text] [doi: 10.1080/24740527.2024.2302487] [Medline: 38482505]
33. Marier-Deschênes P, Pinard AM, Soucy P, Le Capitaine C, Fauteux C, Gagnon-Hamel E. Agir pour moi. Réseau québecois de recherche sur la douleur. URL: https://www.gerermadouleur.ca/agir-pour-moi/ [accessed 2024-04-05]
34. Gandy M, Pang ST, Scott AJ, Heriseanu AI, Bisby MA, Dudeney J, et al. Internet-delivered cognitive and behavioural based interventions for adults with chronic pain: a systematic review and meta-analysis of randomized controlled trials. Pain 2022 Oct 01;163(10):e1041-e1053. [doi: 10.1097/j.pain.0000000000002606] [Medline: 35121696]
35. Michie S, Wood CE, Johnston M, Abraham C, Francis JJ, Hardeman W. Behaviour change techniques: the development and evaluation of a taxonomic method for reporting and describing behaviour change interventions (a suite of five studies involving consensus methods, randomised controlled trials and analysis of qualitative data). Health Technol Assess 2015 Nov;19(99):1-188 [FREE Full text] [doi: 10.3310/hta19990] [Medline: 26616119]
36. Arain M, Campbell MJ, Cooper CL, Lancaster GA. What is a pilot or feasibility study? A review of current practice and editorial policy. BMC Med Res Methodol 2010 Jul 16;10:67 [FREE Full text] [doi: 10.1186/1471-2288-10-67] [Medline: 20637084]
37. Micoulaud-Franchi J, Sauteraud A, Olive J, Sagaspe P, Bioulac S, Philip P. Validation of the French version of the acceptability E-scale (AES) for mental E-health systems. Psychiatry Res 2016 Mar 30;237:196-200. [doi: 10.1016/j.psychres.2016.01.043] [Medline: 26809367]
38. Tariman JD, Berry DL, Halpenny B, Wolpin S, Schepp K. Validation and testing of the acceptability e-scale for web-based patient-reported outcomes in cancer care. Appl Nurs Res 2011 Feb;24(1):53-58 [FREE Full text] [doi: 10.1016/j.apnr.2009.04.003] [Medline: 20974066]
39. Dworkin RH, Turk DC, Farrar JT, Haythornthwaite JA, Jensen MP, Katz NP, IMMPACT. Core outcome measures for chronic pain clinical trials: IMMPACT recommendations. Pain 2005 Jan;113(1-2):9-19. [doi: 10.1016/j.pain.2004.09.012] [Medline: 15621359]
40. Tonkin L. The pain self-efficacy questionnaire. Aust J Physiother 2008;54(1):77. [doi: 10.1016/s0004-9514(08)70073-4] [Medline: 18567187]
41. Nicholas MK. The pain self-efficacy questionnaire: taking pain into account. Eur J Pain 2007 Feb;11(2):153-163. [doi: 10.1016/j.ejpain.2005.12.008] [Medline: 16446108]

XSL•FO
RenderX

42. Poundja J, Fikretoglu D, Guay S, Brunet A. Validation of the French version of the brief pain inventory in Canadian veterans suffering from traumatic stress. J Pain Symptom Manage 2007 Jun;33(6):720-726 [FREE Full text] [doi: 10.1016/j.jpainsymman.2006.09.031] [Medline: 17531912]
43. Tan G, Jensen MP, Thornby JI, Shanti BF. Validation of the brief pain inventory for chronic nonmalignant pain. J Pain 2004 Mar;5(2):133-137 [FREE Full text] [doi: 10.1016/j.jpain.2003.12.005] [Medline: 15042521]
44. Roberge P, Doré I, Menear M, Chartrand E, Ciampi A, Duhoux A, et al. A psychometric evaluation of the French Canadian version of the hospital anxiety and depression scale in a large primary care population. J Affect Disord 2013 May;147(1-3):171-179. [doi: 10.1016/j.jad.2012.10.029] [Medline: 23218249]
45. French DJ, Noël M, Vigneau F, French JA, Cyr CP, Evans RT. L'Échelle de dramatisation face à la douleur PCS-CF: adaptation canadienne en langue française de l'échelle «pain catastrophizing scale». Can J Behav Sci 2005 Jul;37(3):181-192. [doi: 10.1037/h0087255]
46. R Core Team. R: a language and environment for statistical computing. R Foundation for Statistical Computing. URL: https://www.r-project.org/ [accessed 2024-04-05]
47. RStudio: integrated development environment for R. Posit Software. URL: https://posit.co/download/rstudio-desktop/ [accessed 2024-04-05]
48. Bates D, Mächler M, Bolker B, Walker S. Fitting linear mixed-effects models using lme4. J Stat Soft 2015;67(1):1-48. [doi: 10.18637/jss.v067.i01]
49. Lüdecke D, Ben-Shachar M, Patil I, Waggoner P, Makowski D. performance: an R package for assessment, comparison and testing of statistical models. J Open Source Softw 2021 Apr;6(60):3139. [doi: 10.21105/joss.03139]
50. Saunders B, Sim J, Kingstone T, Baker S, Waterfield J, Bartlam B, et al. Saturation in qualitative research: exploring its conceptualization and operationalization. Qual Quant 2018;52(4):1893-1907 [FREE Full text] [doi: 10.1007/s11135-017-0574-8] [Medline: 29937585]
51. Jensen MP, Nielson WR, Kerns RD. Toward the development of a motivational model of pain self-management. J Pain 2003 Nov;4(9):477-492 [FREE Full text] [doi: 10.1016/s1526-5900(03)00779-x] [Medline: 14636816]
52. Creswell JW, Plano Clark VL. Designing and Conducting Mixed Methods Research 3rd edition. Thousand Oaks, CA: SAGE Publications; 2018.
53. Fetters MD, Curry LA, Creswell JW. Achieving integration in mixed methods designs-principles and practices. Health Serv Res 2013 Dec;48(6 Pt 2):2134-2156 [FREE Full text] [doi: 10.1111/1475-6773.12117] [Medline: 24279835]
54. Guetterman TC, Fetters MD, Creswell JW. Integrating quantitative and qualitative results in health science mixed methods research through joint displays. Ann Fam Med 2015 Nov;13(6):554-561 [FREE Full text] [doi: 10.1370/afm.1865] [Medline: 26553895]
55. Yardley L, Spring BJ, Riper H, Morrison LG, Crane DH, Curtis K, et al. Understanding and promoting effective engagement with digital behavior change interventions. Am J Prev Med 2016 Nov;51(5):833-842. [doi: 10.1016/j.amepre.2016.06.015] [Medline: 27745683]
56. Short CE, DeSmet A, Woods C, Williams SL, Maher C, Middelweerd A, et al. Measuring engagement in eHealth and mHealth behavior change interventions: viewpoint of methodologies. J Med Internet Res 2018 Nov 16;20(11):e292 [FREE Full text] [doi: 10.2196/jmir.9397] [Medline: 30446482]
57. Thiblin E, Woodford J, Öhman M, von Essen L. The effect of personalised versus non-personalised study invitations on recruitment within the ENGAGE feasibility trial: an embedded randomised controlled recruitment trial. BMC Med Res Methodol 2022 Mar 06;22(1):65 [FREE Full text] [doi: 10.1186/s12874-022-01553-5] [Medline: 35249543]
58. Meyerowitz-Katz G, Ravi S, Arnolda L, Feng X, Maberly G, Astell-Burt T. Rates of attrition and dropout in app-based interventions for chronic disease: systematic review and meta-analysis. J Med Internet Res 2020 Sep 29;22(9):e20283 [FREE Full text] [doi: 10.2196/20283] [Medline: 32990635]
59. Schlicker S, Baumeister H, Buntrock C, Sander L, Paganini S, Lin J, et al. A web- and mobile-based intervention for comorbid, recurrent depression in patients with chronic back pain on sick leave (Get.Back): pilot randomized controlled trial on feasibility, user satisfaction, and effectiveness. JMIR Ment Health 2020 Apr 15;7(4):e16398 [FREE Full text] [doi: 10.2196/16398] [Medline: 32293577]
60. Guarino H, Fong C, Marsch LA, Acosta MC, Syckes C, Moore SK, et al. Web-based cognitive behavior therapy for chronic pain patients with aberrant drug-related behavior: outcomes from a randomized controlled trial. Pain Med 2018 Dec 01;19(12):2423-2437 [FREE Full text] [doi: 10.1093/pm/pnx334] [Medline: 29346579]
61. Perry J, VanDenKerkhof EG, Wilson R, Tripp DA. Guided internet-based psycho-educational intervention using cognitive behavioral therapy and self-management for individuals with chronic pain: a feasibility study. Pain Manag Nurs 2017 Jun;18(3):179-189. [doi: 10.1016/j.pmn.2016.12.003] [Medline: 28433488]
62. Schultz R, Smith J, Newby JM, Gardner T, Shiner CT, Andrews G, et al. Pilot trial of the reboot online program: an internet-delivered, multidisciplinary pain management program for chronic pain. Pain Res Manag 2018;2018:9634727 [FREE Full text] [doi: 10.1155/2018/9634727] [Medline: 30254712]
63. Carpenter KM, Stoner SA, Mundt JM, Stoelb B. An online self-help CBT intervention for chronic lower back pain. Clin J Pain 2012 Jan;28(1):14-22 [FREE Full text] [doi: 10.1097/AJP.0b013e31822363db] [Medline: 21681084]

XSL•FO
RenderX

64. Rini SC, Porter LS, Somers TJ, McKee DC, DeVellis RF, Smith M, et al. Automated internet-based pain coping skills training to manage osteoarthritis pain: a randomized controlled trial. Pain 2015 May;156(5):837-848 [FREE Full text] [doi: 10.1097/j.pain.0000000000000121] [Medline: 25734997]
65. Chiauzzi E, Pujol LA, Wood M, Bond K, Black R, Yiu E, et al. painACTION-back pain: a self-management website for people with chronic back pain. Pain Med 2010 Jul;11(7):1044-1058. [doi: 10.1111/j.1526-4637.2010.00879.x] [Medline: 20545873]
66. Menga G, Ing S, Khan O, Dupre B, Dornelles AC, Alarakhia A, et al. Fibromyalgia: can online cognitive behavioral therapy help? Ochsner J 2014;14(3):343-349 [FREE Full text] [Medline: 25249800]
67. Mehlsen M, Heegaard L, Frostholm L. A prospective evaluation of the chronic pain self-management programme in a Danish population of chronic pain patients. Patient Educ Couns 2015 May;98(5):677-680. [doi: 10.1016/j.pec.2015.01.008] [Medline: 25662621]
68. Perry J, VanDenKerkhof EG, Wilson R, Tripp DA. Development of a guided internet-based psycho-education intervention using cognitive behavioral therapy and self-management for individuals with chronic pain. Pain Manag Nurs 2017 Apr;18(2):90-101. [doi: 10.1016/j.pmn.2016.12.004] [Medline: 28262481]
69. Ventura J, Sobczak J, Chung J. The chronic knee pain program: a self-management model. Int J Orthop Trauma Nurs 2018 May;29:10-15. [doi: 10.1016/j.ijotn.2018.01.001] [Medline: 29409721]
70. Nicholl BI, Sandal LF, Stochkendahl MJ, McCallum M, Suresh N, Vasseljen O, et al. Digital support interventions for the self-management of low back pain: a systematic review. J Med Internet Res 2017 May 21;19(5):e179 [FREE Full text] [doi: 10.2196/jmir.7290] [Medline: 28550009]
71. Williams DA, Kuper D, Segar M, Mohan N, Sheth M, Clauw DJ. Internet-enhanced management of fibromyalgia: a randomized controlled trial. Pain 2010 Dec;151(3):694-702 [FREE Full text] [doi: 10.1016/j.pain.2010.08.034] [Medline: 20855168]
72. Trudeau KJ, Pujol LA, DasMahapatra P, Wall R, Black RA, Zacharoff K. A randomized controlled trial of an online self-management program for adults with arthritis pain. J Behav Med 2015 Jun;38(3):483-496 [FREE Full text] [doi: 10.1007/s10865-015-9622-9] [Medline: 25686538]

Abbreviations

APM: Agir pour moi
CBT: cognitive behavioral therapy
CHU: University Hospital Centre
CONSORT-EHEALTH: Consolidated Standards of Reporting Trials of Electronic and Mobile Health Applications and Online Telehealth
REDCap: Research Electronic Data Capture

Edited by A Kushniruk; submitted 26.07.23; peer-reviewed by P Gazerani, J Pearson, J Dudeney; comments to author 11.10.23; revised version received 23.10.23; accepted 20.03.24; published 03.05.24.

Please cite as:
Marier-Deschenes P, Pinard AM, Jalbert L, LeBlanc A
Assessing the Feasibility and Preliminary Effects of a Web-Based Self-Management Program for Chronic Noncancer Pain: Mixed Methods Study
JMIR Hum Factors 2024;11:e50747
URL: https://humanfactors.jmir.org/2024/1/e50747
doi:10.2196/50747
PMID:38701440

©Pascale Marier-Deschenes, Anne Marie Pinard, Laura Jalbert, Annie LeBlanc. Originally published in JMIR Human Factors (https://humanfactors.jmir.org), 03.05.2024. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on https://humanfactors.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

A Web-Based Peer Support Network to Help Care Partners of People With Serious Illness: Co-Design Study

Elizabeth A O'Donnell[1,2*], MPH; Aricca D Van Citters[2*], MS; Inas S Khayal[2,3], PhD; Matthew M Wilson[4,5], MD; David Gustafson[6], PhD; Amber E Barnato[2,4,5], MD, MPH, MS; Andrea C Buccellato[3], BFA; Colleen Young[7], BA; Megan M Holthoff[2], MSHS; Eugene Korsunskiy[8], MFA; Stephanie C Tomlin[9], MS; Amelia M Cullinan[4,5], MD; Alexandra C Steinbaugh[5], MSW; Jennifer J Hinson[9], BHCA; Kristen R Johnson[5], MS; Andrew Williams[9], BS; Ruth M Thomson[4,5], DO; Janet M Haines[9], BA; Anne B Holmes[9], BA; Ann D Bradley[9], RN, BSN, MEd; Eugene C Nelson[2], DSC, MPH; Kathryn B Kirkland[2,4,5], MD

[1]Communications, Marketing and Community Health, Alice Peck Day Memorial Hospital, Lebanon, NH, United States
[2]The Dartmouth Institute for Health Policy & Clinical Practice, Geisel School of Medicine at Dartmouth, Lebanon, NH, United States
[3]Dartmouth Cancer Center, Geisel School of Medicine, Dartmouth, Lebanon, NH, United States
[4]Palliative Medicine, Geisel School of Medicine at Dartmouth, Lebanon, NH, United States
[5]Section of Palliative Medicine, Dartmouth Health, Lebanon, NH, United States
[6]College of Engineering, University of Wisconsin, Madison, WI, United States
[7]Mayo Clinic Connect, Mayo Clinic, Rochester, MN, United States
[8]Thayer School of Engineering, Dartmouth College, Hanover, NH, United States
[9]Patient and Family Advisors, Dartmouth Health, Lebanon, NH, United States
[*]these authors contributed equally

Corresponding Author:
Aricca D Van Citters, MS
The Dartmouth Institute for Health Policy & Clinical Practice
Geisel School of Medicine at Dartmouth
One Medical Center Drive, WTRB Level 5
Lebanon, NH, 03756
United States
Phone: 1 603 646 5668
Email: aricca.d.van.citters@dartmouth.edu

Abstract

Background: Care partners of people with serious illness experience significant challenges and unmet needs during the patient's treatment period and after their death. Learning from others with shared experiences can be valuable, but opportunities are not consistently available.

Objective: This study aims to design and prototype a regional, facilitated, and web-based peer support network to help active and bereaved care partners of persons with serious illness be better prepared to cope with the surprises that arise during serious illness and in bereavement.

Methods: An 18-member co-design team included active care partners and those in bereavement, people who had experienced serious illness, regional health care and support partners, and clinicians. It was guided by facilitators and peer network subject-matter experts. We conducted design exercises to identify the functions and specifications of a peer support network. Co-design members independently prioritized network specifications, which were incorporated into an early iteration of the web-based network.

Results: The team prioritized two functions: (1) connecting care partners to information and (2) facilitating emotional support. The design process generated 24 potential network specifications to support these functions. The highest priorities included providing a supportive and respectful community; connecting people to trusted resources; reducing barriers to asking for help; and providing frequently asked questions and responses. The network platform had to be simple and intuitive, provide technical support for users, protect member privacy, provide publicly available information and a private discussion forum, and be easily accessible. It was feasible to enroll members in the ConnectShareCare web-based network over a 3-month period.

XSL•FO
RenderX

Conclusions: A co-design process supported the identification of critical features of a peer support network for care partners of people with serious illnesses in a rural setting, as well as initial testing and use. Further testing is underway to assess the long-term viability and impact of the network.

(JMIR Hum Factors 2024;11:e53194) doi:10.2196/53194

KEYWORDS

human-centered design; caregivers; care partners; serious illness; peer support; online support network; virtual network; online network; caregiver; unmet need; unmet needs; active care; bereaved care; bereavement; clinician; clinicians; function; functions; specification; information; emotional support; technical support; privacy protection; rural; viability; impact; engineering design; care provider; care providers; mortality; quality of life; tertiary care; caregiving

Introduction

Care partners of people with serious illnesses are often overlooked and poorly understood by health care professionals, lack support and educational resources, and are likely to experience significant challenges and unmet needs [1,2]. For many, the work of caring for a person with a serious illness can bring deep satisfaction and can also be challenging [3]. Care partners experience burdens in every area of their lives—emotional, physical, social, spiritual, and financial [2,4,5]. The death of someone with a serious illness (as well as the events leading up to it) also brings hardship, including stress related to loneliness, grief, trauma, role recognition, and self-identity [6-9]. Social isolation and grief are strongly correlated with subsequent depression and related symptoms in bereaved spouses, including sadness, appetite loss, and lower quality of life [5,10]. Bereaved care partners may also face challenges navigating household, financial, social planning, and legal affairs, as well as reintegrating into their local community [11] and accessing available resources. Addressing care partner needs has become a pressing health, economic, and social imperative [12].

Increasingly, care partners are joining online peer support networks to obtain emotional support, access information, and connect and share with others in similar circumstances [13-17]. While people often prefer connecting with health professionals for medical information, they prefer connections with peers over professionals for accessing emotional support or practical advice [18]. In the case of serious illness, when patients may not be well enough to use online peer support networks themselves, care partners are more likely to participate [19].

Previous research has demonstrated a number of variables that contribute to an online peer support network's success or failure [20-22]. Communities that have a clear, defined purpose; foster a strong sense of community; and have a high level of activity are more likely to be successful [20]. Additionally, sustained organizational and financial support for maintaining an online community from inception to maturity is essential, including support for a community manager who sets the tone for the community, creates content, conducts outreach, and fosters a sense of community [20,23]. Successful online networks also harness the interests and abilities of their users to strengthen the community. Networks with more active users are generally more successful [21] because they maintain a critical mass to allow for diversity in experiences and individual attributes, allowing for the natural formation of relationships and answering questions.

Evidence on the impact of online peer support networks for care partners is promising [14,24-26]. For care partners of people with cancer, studies show evidence of decreased care partner emotional distress [27], negative mood [28,29], and sense of burden [29], as well as increases in quality of life and self-efficacy [27]. For care partners of people with dementia, online networks can lead to improvement in self-efficacy [30], decision-making confidence [31], and care partner and patient relationship quality [32]. Care partners also benefit from being able to freely express their sentiments and provide mutual support in a dedicated digital space apart from their loved ones [33,34]. Even people who observe network activity without participating report that reading about the experiences of others is empowering and informative [35,36]. Online networks offer certain advantages: 24/7 home access, flexibility (communication is often asynchronous), anonymity, and a wide range of expertise and experience not limited by geography [37-39].

Online peer support networks for care partners often target specific health conditions (eg, breast cancer and Alzheimer dementia) or stages of caregiving (treatment vs bereavement), but infrequently support care partners of people with diverse conditions or the transition between stages of caregiving. They may also fail to provide active facilitation and moderation; identify and vet regional resources and support from local peers; or provide the possibility of meeting in person. A co-design process can elicit those factors that matter most to the people for whom the network is intended to serve and ensure the successful adoption of the proposed solution [40].

The objective of this paper is to describe a co-design process and the resulting key functions and specifications for a regional, facilitated, and web-based peer support network that can meet the needs of active and bereaved care partners of persons with serious illnesses.

Methods

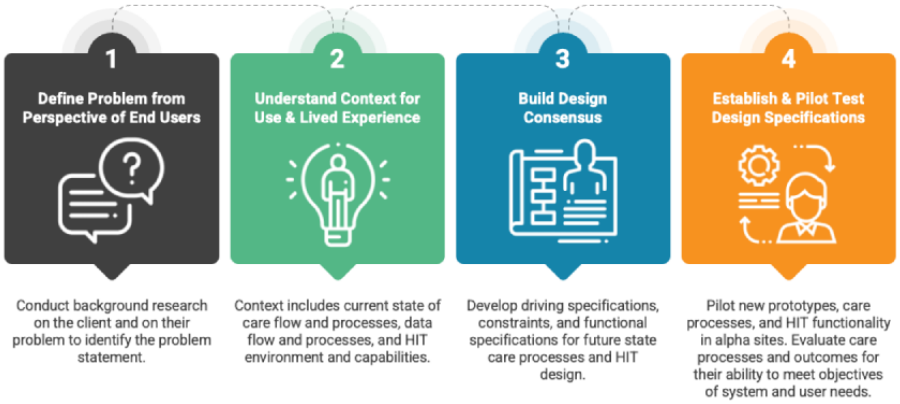
Overview

We applied a co-design framework (Figure 1 [41,42]), which combines human-centered design [43] and engineering design [44,45] processes, to create the specifications for a regional, facilitated, and web-based peer support network. The framework includes 4 stages: defining the problem, understanding the context for use, developing and building consensus around

XSL•FO
RenderX

functions and specifications that fulfill identified needs, and establishing and pilot-testing design specifications.

Figure 1. Co-design framework (reproduced from The Dartmouth Institute for Health Policy & Clinical Practice, which is published under Creative Commons Attribution 4.0 International License). HIT: health information technology.



Ethical Considerations

The study was approved by the Dartmouth College institutional review board (#2000907).

A waiver of written informed consent was used for the surveys, as the only link to the survey respondent would have been the written informed consent document. No identifiable information was collected and individuals were not paid for participating in human subjects research.

Target Population

The target population, hereafter referred to as end users, was defined as care partners (ie, informal caregivers or family members) supporting or providing care to adults (aged 18 years or older) with a serious illness and those who have experienced the loss of someone to a serious illness. The term care partner was chosen by the co-design team to reflect their role and relationship in partnering with a person with a serious illness. Serious illness has been defined as one that carries a high risk of mortality and either negatively impacts a person's daily function or quality of life, or excessively strains their care partners [46]. Care partners include relatives, spouses or partners, friends, neighbors, or others who have a significant personal relationship with, and who provide a broad range of assistance to, a person with a serious illness. We focused on care partners living in New Hampshire and Vermont, the catchment area for Dartmouth-Hitchcock Medical Center.

Participants

We formed an 18-member team to co-design the network. The team included 2 active and bereaved care partners of people with serious illness, 4 adults with serious illness, 6 interdisciplinary palliative care clinical team members, and 6 support service staff. Care partners and people with serious illnesses were recruited by our clinical team partners. Clinical team members were affiliated with Dartmouth-Hitchcock Medical Center, a rural tertiary care academic medical center in New Hampshire. Facilitation of the co-design process was led by researchers with expertise in co-design, evaluation, and quality improvement (EAO and ADVC). The design process was informed by consultation with an expert in human-centered design (EK) and a systems engineer (ISK). To ensure our design aligned with best practices, we consulted with external advisors with expertise in facilitated support networks (DG and CY), met with regional health care and support partners, and obtained input from an external advisory committee with expertise in scaling innovations, business, and serious illness.

The co-design team met twice a month for 8 months (April to November 2019) to identify and prioritize the functions and specifications of the network and met monthly for 6 months (December 2019 to May 2020) to test prototypes.

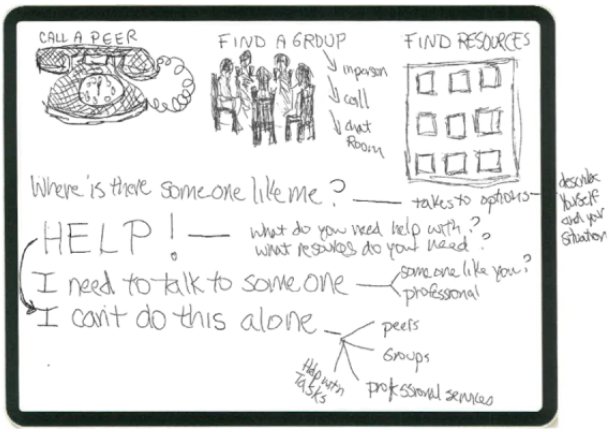
Defining the Problem and Understanding Context for Use

We conducted human-centered design exercises [43] to elicit community needs and assets, define the problem, and understand the context for use. We drew upon stories of serious illnesses shared by care partners to identify needs arising from lived experiences. Design facilitators shadowed [43] outpatient palliative care visits and attended interdisciplinary palliative care team meetings to further understand the context of use; services provided by the care team; and the daily lives of people with serious illness, their care partners, and clinicians [47]. We developed empathy maps [43] to reflect and articulate what end users hear, see, say, do, think, and feel, and to identify points of pain and gain. We used a visual thinking exercise [48] to sketch ideas for an ideal support network (example, Figure 2). Design exercises were reviewed during design sessions to discuss critical functions and important features that fulfill and support these functions.

We supplemented design activities with surveys of potential end users. Between December 2018 and April 2019, we collected 28 surveys from a convenience sample of active care partners presenting at the Dartmouth-Hitchcock Medical Center outpatient palliative care clinic and 21 surveys from a convenience sample of bereaved care partners affiliated with the clinic or regional health organizations (eg, hospice) to elicit information on the challenges, needs, and desire for peer connection among active and bereaved care partners (surveys are provided in Multimedia Appendix 1). We used descriptive statistics to summarize categorical data, and thematic analyses to identify themes from open-ended questions.

XSL•FO
RenderX

Figure 2. Sketch of a serious illness community network, created by a co-design team member.



Building Design Consensus

We collated, organized, and systematically described the identified needs into a hierarchical list. We converted the highest-level needs into functions for the network to achieve. At the highest level, for example, "care partner need for information" was converted to "connecting care partners to information" and "care partner need for emotional support" was converted to "facilitating emotional support." We identified potential specifications associated with each function of the network.

Co-design team members independently rated the importance of each specification using a five-item prioritization scale: (5) must have, (4) high importance (feasible without), (3) should have (very important), (2) could have (consideration), and (1) desirable (will not have at this time) [49]. Team members were provided an opportunity to describe their understanding and thought process for preference identification as a group. Scores were weighted and averaged by respondent type to ensure that patients and care partners, clinicians, and support service staff had equal participation weight (eg, scores from 6 clinicians were averaged to create 1 clinician average).

Pilot-Testing Design Specifications

We tested a series of prototypes to identify the importance of different network components. These included storytelling exercises, face-to-face group conversations, one-to-one matching of care partners, group videoconferencing, an online discussion forum, and a "Caregiver Day" event at the health center. Summaries are provided in Multimedia Appendix 2.

After finalizing design functions and specifications, we identified potential vendors to host the network by conducting an environmental scan of web-based networks focused on people with serious illnesses and care partners and mapping desired design specifications to vendor capabilities. Four vendors provided demonstrations of their platforms between April and May 2020. Vendor selection was driven by the vendor's ability to provide the prioritized functions and specifications, the cost to build and maintain the network, and being a US-based company. Following vendor selection, the web-based network platform was customized by the design team to deliver upon design functions and specifications.

User acceptance testing was conducted between April and July 2021. Four care partners and 2 people who previously had a serious illness were invited to register as founding members of the network in April. These founding members were encouraged to invite care partners they knew into the web-based community to test the feasibility of enrolling members. Three clinical champions (physicians and social workers), a chaplain, and a staff member who manages complementary care programs referred care partners to the network. The research team met weekly with the vendor during the user acceptance testing period to resolve issues. The web-based platform remained available for registered members to use while issues were addressed. The network moderator met monthly with the design team to plan future improvements using a quality improvement framework [50].

Results

Problem Definition

The co-design process led to clarity around the objective of the network: to help care partners cope with the surprises that arise during serious illness and bereavement. The network, named "ConnectShareCare," was intended to supplement existing services, to be provided outside of clinical encounters with the health care system or regional professional support and service organizations, and to tap into the wisdom of those with lived experiences.

If successful, the design team anticipated that the network would benefit 4 groups, as outlined in Table 1.

XSL•FO
RenderX

Table 1. Anticipated impact of ConnectShareCare.

Audience	Anticipated objectives
Care partners	• Improve access to information that can guide decision-making • Improve sense of empowerment in making decisions and providing support • Decrease sense of distress and social isolation
Community partners	• Improve understanding of needs and gaps in service • Provide a system to share assets or resources
Clinicians or health care system	• Address gaps in services that are not currently met • Improve availability to see patients who seek services • Improve efficiency of health care encounters
Quality improvement leaders and researchers	• Improve understanding of the needs of care partners • Align services with care partners and community needs • Demonstrate a positive impact of the network over time

Context for Use and Lived Experiences

Our co-design process identified that active and bereaved care partners have different needs but have common interests in sharing information and providing or receiving support. Active care partners who completed an assessment survey were most challenged by emotional difficulties (eg, worry, uncertainty, or lack of control; 12/28, 43%), providing care and emotional support (7/28, 25%), and practical matters (6/28, 21%). Bereaved care partners were most challenged by loneliness (10/21, 48%), managing grief and emotional difficulties (6/21, 29%), and managing practical matters (5/21, 24%). Active care partners were most helped by support from friends, family, or other social connections (12/28, 43%), as well as by medical professionals (9/28, 32%), while bereaved care partners were most helped by support from friends, family, or other social connections (16/21, 76%) and by developing self-care strategies that led to personal resilience and growth (12/21, 57%).

Most active (18/27, 67%) and bereaved (18/21, 86%) care partners were interested in 1 or more forms of connecting with other people who have shared a similar care experience. Both active and bereaved care partners anticipated that a network could provide support, knowledge, and resources but anticipated challenges associated with time to participate and with forming personal connections.

The series of human-centered design exercises and interviews led to additional insights. First, active and bereaved care partners may benefit from connecting and sharing information with each other. Peer-generated information from care partners who have shared a similar experience feels more authentic, detailed, and actionable. Second, care partners wished to belong to a local support community that was connected through geographic proximity and could provide recommendations for local resources. Third, a web-based network enables care partners to access information at any place or time, allows anonymity, improves access for people who are home-bound or grieving, and may reach an increased number of care partners. Fourth, trained staff who can moderate, promote, and manage the web-based community and volunteers who can recruit and engage users are important. Paid or volunteer moderators can play an important role in listening, making connections, and highlighting information, services, and programs. Fifth, a web-based network would benefit from supplemental opportunities for the community to meet face-to-face or through digital programming.

The design process also identified several potential risks and possible mitigation strategies. First, there was a risk of causing harm to vulnerable end users if the design failed to provide a safe and supportive environment, protect the privacy of sensitive information, or enact acceptable data ownership guidelines. Second, there was a risk associated with the usability of the network among end users who were less facile with web-based services. Third, there was a risk associated with the inability to form a personal connection with peers through a web-based network. Other potential risks included those associated with competition from other networks, inaccurate content or information, and care partners having minimal time to participate in a web-based support network due to other responsibilities.

Build Design Consensus

Two primary functions emerged from the design activities: to support care partners in (1) providing each other with emotional support and (2) exchanging helpful information and resources. We developed specifications related to these functions (Table 2), as well as the form of the network, including the user interface, data and security, and other considerations (Table 3).

XSL•FO
RenderX

Table 2. Prioritization ranking of design specifications associated with network functions (weighted average scores across patients and families [n=5]; clinicians [n=4]; clinical support service providers [n=5]; and network advisors [n=2]).

Network functions and prioritization ranking of associated design specifications	Ranking[a]
Function 1: provide each other with emotional support	
Must have	
• Provide a supportive and respectful space • Ability for established "Guidelines and Ground Rules" to be clearly visible to users • Ability to protect individual identity (opt-out options for sharing personal information; opportunity to keep geographic location private) • Allows moderator functionality for policing interactions and blocking users if necessary	5.00
Should have (very important)	
• Incorporate and help facilitate, one-to-one connections • Ability to locate "true peer" (similar users) through the platform via matching on similar life circumstances (through backend algorithm or user profile details: type of loss, disease, time caregiving, or time since loss) • Allow for private one-to-one messaging to facilitate a more personal connection, not monitored by an external entity	3.27
• Includes opportunity for storytelling based on personal user content or experience	3.00
• Provides an opportunity to share solutions	3.00
Could have (consideration)	
• Differentiate between whether people want to feel heard or want to hear solutions • Ability for users to designate whether they are looking to hear solutions or feedback or simply share	2.88
Function 2: exchange helpful information and resources	
Must have	
• Provide connections to trusted and curated local, national, and international resources • Ability to host webinars in order to share educational content • Ability for newsfeed or wall that features newly published content • Provide document or resource repository related to user needs • Ability to highlight and share local events via calendar, bulletin board, or other • Robust search function available to find targeted resources within the platform	4.56
High importance (feasible without)	
• Reduce the difficulty of asking for help by normalizing needing help • Ability to add a button in various locations that asks "Having a hard time asking for help?" and that opens a new page that contains tips or guidelines on how to ask for help and what to expect when asking for help	3.88
• Provide frequently asked questions list and answers • Site provides a list, or ability to create a list, of the most popular or frequently asked questions and answers (eg, "How do I cope with stress?")	3.76
Should have (very important)	
• Ability to identify most common needs • Ability to organize conversations around themes (or topics) and make it easy for someone with a specific question, topic, or theme to locate information pertaining to it • Ability to "like," (showing interest, support) posts, topics, or comments so that users can see which posts are popular and most useful • Ability to follow a discussion thread, topic, etc. Once a user has "followed" something or someone, they can receive a notification when there is new content posted • Ability to bookmark posts (to save content) that users would like to revisit	3.47
• Ability to be supported by local or regional expert moderator (community manager)	3.38

XSL•FO
RenderX

Network functions and prioritization ranking of associated design specifications	Ranking[a]
• Include the ability to publish videos related to the content of the network • Allow users to record and post videos instantly (personal and other)	3.06

[a]Prioritization ranking: (5) must have, (4) high priority (feasible without), (3) very important (should have), (2) consideration (could have), (1) desirable (will not have at this time).

XSL•FO
RenderX

Table 3. Prioritization ranking of the form of the network: user interface, data and security, and other considerations (weighted average scores across patients and families [n=5]; clinicians [n=4]; clinical support service providers [n=5]; and network advisors [n=2]).

Form of the network and its prioritization ranking	Ranking[a]
User interface	
Must have	
• Simple or intuitive interface • Provide support, guidance, and assistance with how to navigate and use platform (ideal: offer video tutorials) • Passes Web Content Accessibility Guidelines (WCAG). Example: large font • Ability to easily identify new content (since user's last login)	5.00
• Provides IT technical support (for members)	5.00
• Easy access to support user engagement • Smooth and simple login process • Optimized for mobile device • Real-time information and comments available and accessible (not prescreened by community moderator)	4.53
High importance (feasible without)	
• Aesthetically refined • Pleasing to the eye, organized, and appropriate imagery • Symmetrical and aligned (looks modern)	4.00
• Appropriate use of pop-ups and other interactive elements	4.00
• Does not allow advertising • Does not have advertisements on the platform itself • Does not send any unsolicited promotional emails related to the platform or other	3.76
Data and security	
Must have	
• Secure platform or user privacy protected (Health Insurance Portability and Accountability Act [HIPAA] compliant)	5.00
• Public forum for information and resources, but opportunities for private discussion forums	5.00
Should have (very important)	
• Data are owned by the co-design team's institution (not the vendor) • Establish terms and conditions for how information or data will be accessed, stored, and used • No selling of data to for-profit or not-for-profit entities (pharmaceuticals or other) for financial gains	3.47
• Data access and analysis • Provides actionable metrics related to user activity and engagement (including IP address) • Data analysis capabilities available within local database • Create an extract of selected data • Ability to survey users	3.00
Other considerations	
High importance (feasible without)	
• Health-focused support network • Network software is targeted to health-focused communities, has features and functions relevant to health, self-management, etc (eg, health needs assessment) and has experience working in peer-to-peer health care	4.00
• Sustainability (retaining users) • Provides facilitation and network growth support (through designated pump primers, marketing, etc) • Opportunity to offer member incentives (through badges, quality improvement initiatives: creating educational material, etc)	4.00

XSL•FO
RenderX

Form of the network and its prioritization ranking	Ranking[a]
• Scalability • Interactive and responsive; ability to customize and add measures and functionality over time	4.00

[a]Prioritization ranking: (5) must have, (4) high priority (feasible without), (3) very important (should have), (2) consideration (could have), (1) desirable (will not have at this time).

The most highly prioritized specifications to support each function (Table 2) included providing a supportive and respectful space; providing connections to trusted and curated local, national, and international resources; reducing the difficulty of asking for help by normalizing needing help; and providing curated resources to address the most common concerns (eg, easy access to frequently asked questions and answers).

The user interface (Table 3) must be simple and intuitive, provide technical support for users, and be easy to access. It was highly important for the user interface to be aesthetically refined, include appropriate use of interactive elements, and not allow external advertising. The platform must be secure and protect user privacy. It must be available as a public forum for information but also allow participants to communicate via discussions not visible to others outside of the network. Other highly important considerations included hosting by a vendor with experience in providing health-focused networks and providing features that support sustainability and scalability (such as member incentives and the ability to customize or add functionality over time).

Pilot Test Design Specifications

Following the vendor selection process, we worked with CareHubs, an online health network vendor, to build the ConnectShareCare network. The network included (1) a single support community for active and bereaved care partners; (2) a short list of curated resources based on needs identified during the design (planning ahead, practical issues, emotional issues, communication issues, and family resources); (3) a calendar of online and in-person regional support programming; (4) a story from an end user about their experience as a care partner; (5) a roster with member profiles; (6) a help center; and (7) community guidelines. Screenshots are included in Multimedia Appendix 3.

The initial feasibility of the network was demonstrated through the active participation of founding members, beginning in April 2021, with expansion to include 12 new care partner members in May 2021 and 15 new members in June 2021. During this period, the network had an average of 135 posts per month. A total of 16 members posted a message to the network, with an average of 11 members posting per month.

Discussion

Principal Findings

A co-design process generated a useful and feasible regional, facilitated, and web-based peer support network for care partners of people with serious illnesses. The co-design process ensured that all voices were heard, especially among people who typically may not work together. Design decisions were made collectively and systemically, which allowed network functions and specifications to be identified and prioritized. By doing so, the co-design process ensured that the most critical decisions were responsive to regional needs and preferences. The resulting network connects active and bereaved care partners with peers to facilitate emotional support and exchange information related to caregiving for people with serious illness.

Comparison With Prior Work

Similar to other networks [15,33,51-62], ConnectShareCare has a clear purpose, includes mechanisms to foster a strong sense of community and support among regional care partners, and provides value to a variety of groups. The network builds on the resources, wisdom, and experiences of care partners. The inclusion of a moderator helps ensure a safe environment that is protected from misinformation, trolling, or cyberbullying. The moderator sets the tone and etiquette with members, modeling behavior and other preventative measures, as well as moderating posts, facilitating connections, and providing feedback to adjust member behaviors [20,23].

Our design process had several strengths. First, our process engaged people who would be end users of the network in making critical design decisions [40]. In contrast to asking end users about single decisions, our process allowed end users to make decisions in the context of all other design requirements and options. This led to a more systemic approach to engagement that ensures that decisions are optimized to fit together. Second, our process brought together people with different expertise (in being a person with serious illness or a care partner, in medicine, in health network design, and in community management and moderation) that may typically not work together to create services, and each had different needs to maintain the value of the network. The process also engaged people with engineering design [44,45] and user-centered design [43] expertise to ensure that the process was rigorous enough to produce a network tailored to the needs of end users. Third, our design was responsive to regional needs by addressing gaps in available services and drawing upon local assets. We intentionally worked with multiple health care and support organizations in the region to broaden our network and reach care partners most in need of support, regardless of where formal health care services were received. Care partners are a vulnerable population who often receive minimal structured support from the health system, yet they have significant knowledge to contribute on how to navigate health care systems, health and social resources, and losses at every stage (eg, relationships, identity, and freedom). This knowledge is often actionable by peers. Finally, our process may have supported growth in network participation due to health care and regional

XSL•FO
RenderX

support partners feeling heard and included in the design process.

Limitations

This project has certain limitations that should be considered by those wishing to adapt the methods and findings to different contexts. First, the project was built around a recognition that people who live in north-east rural areas can be particularly isolated and lack access to sources of support. It is unclear whether the network will meet the needs of people in other regions of the country. Second, similar to local demographic characteristics, our design team had limited racial or ethnic diversity. We do not know how greater racial, ethnic, or sociodemographic diversity would enhance or create barriers to its success. Third, the network is built around an asynchronous model. This can be very important because it respects the different schedules that people are on; however, it limits the opportunities for people to hear and see each other in real time. It is unknown whether a network that combines synchronous and asynchronous components would be useful in our context. Finally, while the design process requires extensive back and forth among participants and may not be feasible in other situations, it also represents a strength in creating a network that more closely reflects community needs. In our situation, the decision to create a new network was a result of the recognition that other solutions are not likely to fulfill the needs of our end users.

Conclusions

Care partners of people with serious illness often lack support and are likely to experience significant challenges and unmet needs. We followed a structured co-design process to collaboratively identify and prioritize the functions and specifications of a regional web-based facilitated peer support network to help care partners cope with the surprises that arrive during serious illness and bereavement. The network was designed to provide emotional support and exchange information related to serious illness caregiving. The coproduction of accessible peer-led information, resources, and support may extend the scope of services offered by a health system to support lay care partners—becoming part of a sustainable, person-centered value-creation system [63,64]. Opportunities exist to evaluate the feasibility of actively engaging community members and moderators in the network [65,66] and will be reported on in a publication under development. Moreover, there is a need to understand effective mechanisms to recruit and retain participants and provide a safe environment to people who are in vulnerable situations; to monitor the network life cycle through metrics related to activity and growth [20,22]; and to consider the creation of network subgroups to support care partners of people with particular illnesses (eg, cancer, dementia, and Parkinson disease) or people from specific minority populations. Finally, there is an opportunity to understand the impact of a regional support network on care partner quality of life, self-confidence, loneliness, and isolation [67,68]; and on health system reputation, use, and visibility.

Acknowledgments

The authors wish to thank the many other people who contributed to the design and implementation of the ConnectShareCare network and the many active and bereaved care partners who participated in and provided feedback on the network. Special thanks are given to Deadra Ashton, James Bowling, Mary Elizabeth Byrnes, Nancy Duhaine, Ellen Flaherty, Melissa Garland, Malavika Govindan, Amanda Hoggard, Marv Klassen-Landis, Jane Masters, Catherine Reed, and Donna Soltura for their feedback and contributions to the design of the network. In addition, the authors wish to thank the Upper Valley Elder Forum; Upper Valley Bereavement Partnership Group; Dartmouth-Hitchcock: Volunteers, Palliative Care Program, Arts and Humanities Program, and Aging Resource Center; Dartmouth Cancer Center; Visiting Nurse and Hospice of Vermont and New Hampshire; Lake Sunapee Visiting Nurses Association and Hospice Care; Bayada Home Care; CommunityCare of Lyme; Community Nurse Connection; Shire Digital; and CareHubs for their support of the ConnectShareCare network. This work was supported by the Gordon and Betty Moore Foundation (grant number 7485) and the Couch Family Fund at The Dartmouth Institute for Health Policy and Clinical Practice.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Surveys administered to active and bereaved caregivers.
[DOCX File , 479 KB - humanfactors_v11i1e53194_app1.docx]

Multimedia Appendix 2
Prototype and testing results.
[DOCX File , 23 KB - humanfactors_v11i1e53194_app2.docx]

Multimedia Appendix 3
ConnectShareCare screenshots.
[DOCX File , 1751 KB - humanfactors_v11i1e53194_app3.docx]

XSL•FO
RenderX

References

1. Northouse LL, Katapodi MC, Song L, Zhang L, Mood DW. Interventions with family caregivers of cancer patients: meta-analysis of randomized trials. CA Cancer J Clin 2010;60(5):317-339 [FREE Full text] [doi: 10.3322/caac.20081] [Medline: 20709946]
2. Choi S, Seo J. Analysis of caregiver burden in palliative care: an integrated review. Nurs Forum 2019;54(2):280-290. [doi: 10.1111/nuf.12328] [Medline: 30737798]
3. Lewis C, Abrams MK, Schneider EC, Shah T. Coping with serious illness in America: relying on family and friends. The Commonwealth Fund. 2019. URL: https://www.commonwealthfund.org/publications/2019/feb/coping-serious-illness-america [accessed 2024-04-10]
4. McDonald J, Swami N, Pope A, Hales S, Nissim R, Rodin G, et al. Caregiver quality of life in advanced cancer: qualitative results from a trial of early palliative care. Palliat Med 2018;32(1):69-78. [doi: 10.1177/0269216317739806] [Medline: 29130418]
5. Götze H, Brähler E, Gansera L, Schnabel A, Gottschalk-Fleischer A, Köhler N. Anxiety, depression and quality of life in family caregivers of palliative cancer patients during home care and after the patient's death. Eur J Cancer Care (Engl) 2018;27(2):e12606. [doi: 10.1111/ecc.12606] [Medline: 27859889]
6. Utz RL, Swenson KL, Caserta M, Lund D, deVries B. Feeling lonely versus being alone: loneliness and social support among recently bereaved persons. J Gerontol B Psychol Sci Soc Sci 2014;69(1):85-94 [FREE Full text] [doi: 10.1093/geronb/gbt075] [Medline: 24056690]
7. U.S. Surgeon General. Our epidemic of loneliness and isolation: the U.S. surgeon general's advisory on the healing effects of social connection and community. U.S. Department of Health and Human Services. 2023. URL: https://www.hhs.gov/sites/default/files/surgeon-general-social-connection-advisory.pdf [accessed 2024-04-10]
8. Ugalde A, Krishnasamy M, Schofield P. Role recognition and changes to self-identity in family caregivers of people with advanced cancer: a qualitative study. Support Care Cancer 2012;20(6):1175-1181. [doi: 10.1007/s00520-011-1194-9] [Medline: 21611864]
9. Duffy M, Wild J. Living with loss: a cognitive approach to prolonged grief disorder—incorporating complicated, enduring and traumatic grief. Behav Cogn Psychother 2023;51(6):645-658 [FREE Full text] [doi: 10.1017/S1352465822000674] [Medline: 37170765]
10. Fried EI, Bockting C, Arjadi R, Borsboom D, Amshoff M, Cramer AOJ, et al. From loss to loneliness: the relationship between bereavement and depressive symptoms. J Abnorm Psychol 2015;124(2):256-265. [doi: 10.1037/abn0000028] [Medline: 25730514]
11. Kim Y, Carver CS, Cannady RS. Bereaved family cancer caregivers' unmet needs: measure development and validation. Ann Behav Med 2020;54(3):164-175 [FREE Full text] [doi: 10.1093/abm/kaz036] [Medline: 31602463]
12. Reinhard SC, Caldera S, Houser A, Choula RB. Valuing the invaluable 2023 update: strengthening supports for family caregivers. AARP Public Policy Institute. 2023. URL: https://www.aarp.org/ppi/info-2015/valuing-the-invaluable-2015-update.html [accessed 2024-04-10]
13. Fox S, Duggan M, Purcell K. Family caregivers are wired for health. Pew Research Center. Washington, D.C; 2013. URL: https://www.pewresearch.org/internet/2013/06/20/family-caregivers-are-wired-for-health/ [accessed 2024-04-10]
14. Shin JY, Kang TI, Noll RB, Choi SW. Supporting caregivers of patients with cancer: a summary of technology-mediated interventions and future directions. Am Soc Clin Oncol Educ Book 2018;38:838-849 [FREE Full text] [doi: 10.1200/EDBK_201397] [Medline: 30231412]
15. Gage-Bouchard EA, LaValley S, Mollica M, Beaupin LK. Cancer communication on social media: examining how cancer caregivers use facebook for cancer-related communication. Cancer Nurs 2017;40(4):332-338. [doi: 10.1097/NCC.0000000000000418] [Medline: 27442210]
16. Ahmad A, Premanandan S, Langegård U, Cajander Å, Johansson B, Carlsson M, et al. Positive design framework for carer eSupport: qualitative study to support informal caregivers of patients with head and neck cancer in Sweden. JMIR Cancer 2023;9:e45748 [FREE Full text] [doi: 10.2196/45748] [Medline: 37252765]
17. Dale J, Nanton V, Day T, Apenteng P, Bernstein CJ, Grason Smith G, et al. Uptake and use of care companion, a web-based information resource for supporting informal carers of older people: mixed methods study. JMIR Aging 2023;6:e41185 [FREE Full text] [doi: 10.2196/41185] [Medline: 37733406]
18. Fox S. Peer to peer healthcare: many people—especially those living with chronic or rare diseases—use online connections to supplement professional medical advice. Pew Research Center. Washington DC; 2011. URL: https://www.pewresearch.org/internet/2011/02/28/peer-to-peer-health-care-2/ [accessed 2024-04-10]
19. Kinnane NA, Milne DJ. The role of the internet in supporting and informing carers of people with cancer: a literature review. Support Care Cancer 2010;18(9):1123-1136. [doi: 10.1007/s00520-010-0863-4] [Medline: 20336326]
20. Young C. Community management that works: how to build and sustain a thriving online health community. J Med Internet Res 2013;15(6):e119 [FREE Full text] [doi: 10.2196/jmir.2501] [Medline: 23759312]
21. Iriberri A, Leroy G. A life-cycle perspective on online community success. ACM Comput Surv 2009;41(2):1-29. [doi: 10.1145/1459352.1459356]

XSL•FO
RenderX

22. Preece J. Sociability and usability in online communities: determining and measuring success. Behav Inform Technol 2010;20(5):347-356. [doi: 10.1080/01449290110084683]
23. Skousen T, Safadi H, Young C, Karahanna E, Safadi S, Chebib F. Successful moderation in online patient communities: inductive case study. J Med Internet Res 2020;22(3):e15983 [FREE Full text] [doi: 10.2196/15983] [Medline: 32181743]
24. Oliver DP, Patil S, Benson JJ, Gage A, Washington K, Kruse RL, et al. The effect of internet group support for caregivers on social support, self-efficacy, and caregiver burden: a meta-analysis. Telemed J E Health 2017;23(8):621-629. [doi: 10.1089/tmj.2016.0183] [Medline: 28328392]
25. Ploeg J, Ali MU, Markle-Reid M, Valaitis R, Bartholomew A, Fitzpatrick-Lewis D, et al. Caregiver-focused, web-based interventions: systematic review and meta-analysis (part 2). J Med Internet Res 2018;20(10):e11247 [FREE Full text] [doi: 10.2196/11247] [Medline: 30368439]
26. Dam AEH, de Vugt ME, Klinkenberg IPM, Verhey FRJ, van Boxtel MPJ. A systematic review of social support interventions for caregivers of people with dementia: are they doing what they promise? Maturitas 2016;85:117-130. [doi: 10.1016/j.maturitas.2015.12.008] [Medline: 26857890]
27. Northouse L, Schafenacker A, Barr KLC, Katapodi M, Yoon H, Brittain K, et al. A tailored web-based psychoeducational intervention for cancer patients and their family caregivers. Cancer Nurs 2014;37(5):321-330 [FREE Full text] [doi: 10.1097/NCC.0000000000000159] [Medline: 24945270]
28. DuBenske LL, Gustafson DH, Namkoong K, Hawkins RP, Atwood AK, Brown RL, et al. CHESS improves cancer caregivers' burden and mood: results of an eHealth RCT. Health Psychol 2014;33(10):1261-1272 [FREE Full text] [doi: 10.1037/a0034216] [Medline: 24245838]
29. Chih MY, DuBenske LL, Hawkins RP, Brown RL, Dinauer SK, Cleary JF, et al. Communicating advanced cancer patients' symptoms via the internet: a pooled analysis of two randomized trials examining caregiver preparedness, physical burden, and negative mood. Palliat Med 2013;27(6):533-543 [FREE Full text] [doi: 10.1177/0269216312457213] [Medline: 22988042]
30. Marziali E, Garcia LJ. Dementia caregivers' responses to 2 internet-based intervention programs. Am J Alzheimers Dis Other Demen 2011;26(1):36-43 [FREE Full text] [doi: 10.1177/1533317510387586] [Medline: 21282276]
31. Bass DM, McClendon MJ, Brennan PF, McCarthy C. The buffering effect of a computer support network on caregiver strain. J Aging Health 1998;10(1):20-43. [doi: 10.1177/089826439801000102] [Medline: 10182416]
32. McKechnie V, Barker C, Stott J. The effectiveness of an internet support forum for carers of people with dementia: a pre-post cohort study. J Med Internet Res 2014;16(2):e68 [FREE Full text] [doi: 10.2196/jmir.3166] [Medline: 24583789]
33. Oliver DP, Washington K, Wittenberg-Lyles E, Gage A, Mooney M, Demiris G. Lessons learned from a secret Facebook support group. Health Soc Work 2015;40(2):125-133. [doi: 10.1093/hsw/hlv007] [Medline: 26027421]
34. Demiris G, Washington K, Ulrich CM, Popescu M, Oliver DP. Innovative tools to support family caregivers of persons with cancer: the role of information technology. Semin Oncol Nurs 2019;35(4):384-388 [FREE Full text] [doi: 10.1016/j.soncn.2019.06.013] [Medline: 31229338]
35. van Uden-Kraan CF, Drossaert CHC, Taal E, Shaw BR, Seydel ER, van de Laar MAFJ. Empowering processes and outcomes of participation in online support groups for patients with breast cancer, arthritis, or fibromyalgia. Qual Health Res 2008;18(3):405-517. [doi: 10.1177/1049732307313429] [Medline: 18235163]
36. Han JY, Hou J, Kim E, Gustafson DH. Lurking as an active participation process: a longitudinal investigation of engagement with an online cancer support group. Health Commun 2014;29(9):911-923 [FREE Full text] [doi: 10.1080/10410236.2013.816911] [Medline: 24345206]
37. Colvin J, Chenoweth L, Bold M, Harding C. Caregivers of older adults: advantages and disadvantages of internet-based social support*. Fam Relat 2003;53(1):49-57. [doi: 10.1111/j.1741-3729.2004.00008.x]
38. Hopwood J, Walker N, McDonagh L, Rait G, Walters K, Iliffe S, et al. Internet-based interventions aimed at supporting family caregivers of people with dementia: systematic review. J Med Internet Res 2018;20(6):e216 [FREE Full text] [doi: 10.2196/jmir.9548] [Medline: 29895512]
39. Lee E. Do technology-based support groups reduce care burden among dementia caregivers? a review. J Evid Inf Soc Work 2015;12(5):474-487. [doi: 10.1080/15433714.2014.930362] [Medline: 25794367]
40. Sanders EBN, Stappers PJ. Co-creation and the new landscapes of design. CoDesign 2008;4(1):5-18 [FREE Full text] [doi: 10.1080/15710880701875068]
41. Van Citters AD. CC BY 4.0 DEED attribution 4.0 international. Creative Commons. 2023. URL: https://creativecommons.org/licenses/by/4.0/ [accessed 2024-04-10]
42. Tips and tools. The Dartmouth Institute for Health Policy & Clinical Practice. URL: https://sites.dartmouth.edu/coproduction/knowledge-center/tools/ [accessed 2024-04-26]
43. Stickdorn M, Schneider J. This is Service Design Thinking: Basics, Tools, Cases. Amsterdam: BIS Publishers; 2011.
44. Buede DM, Miller WD. The Engineering Design of Systems: Models and Methods. Hoboken, NJ: John Wiley & Sons, Inc; 2016.
45. Khayal IS. Designing technology and healthcare delivery systems to support clinician and patient care experiences: a multi-stakeholder systems engineering co-design methodology. : IEEE; 2019 Presented at: 2019 IEEE International

XSL•FO
RenderX

Symposium on Technology and Society (ISTAS); November 15-16, 2019; Medford, MA, USA p. 1-6. [doi: 10.1109/istas48451.2019.8937932]
46. Kelley AS, Bollens-Lund E. Identifying the population with serious illness: the "denominator" challenge. J Palliat Med 2018;21(S2):S7-S16 [FREE Full text] [doi: 10.1089/jpm.2017.0548] [Medline: 29125784]
47. Khayal IS, Barnato AE. What is in the palliative care 'syringe'? a systems perspective. Soc Sci Med 2022;305:115069 [FREE Full text] [doi: 10.1016/j.socscimed.2022.115069] [Medline: 35691210]
48. Rohde M. Sketching: the visual thinking power tool. A List Apart. 2011. URL: https://alistapart.com/article/sketching-the-visual-thinking-power-tool/ [accessed 2024-04-10]
49. Agile Business Consortium. Chapter 10: MoSCoW prioritisation. DSDM Project Framework. 2014. URL: https://www.agilebusiness.org/dsdm-project-framework/moscow-prioririsation.html [accessed 2024-04-10]
50. Langley GJ, Nolan KM, Nolan TW, Norman CL, Provost LP, Moen RD. The Improvement Guide: A Practical Approach to Enhancing Organizational Performance, 2nd Edition. San Francisco: Jossey-Bass Publishers; 2009.
51. Hwang J, Toma CL, Chen J, Shah DV, Gustafson D, Mares ML. Effects of web-based social connectedness on older adults' depressive symptoms: a two-wave cross-lagged panel study. J Med Internet Res 2021;23(1):e21275 [FREE Full text] [doi: 10.2196/21275] [Medline: 33439143]
52. Gustafson DH, Kornfield R, Mares ML, Johnston DC, Cody OJ, Yang EF, et al. Effect of an eHealth intervention on older adults' quality of life and health-related outcomes: a randomized clinical trial. J Gen Intern Med 2022;37(3):521-530 [FREE Full text] [doi: 10.1007/s11606-021-06888-1] [Medline: 34100234]
53. Kotsenas AL, Aase L, Arce M, Timimi FK, Dacy M, Young C, et al. The social media DNA of Mayo Clinic-and health care. J Am Coll Radiol 2018;15(1 Pt B):162-166 [FREE Full text] [doi: 10.1016/j.jacr.2017.09.026] [Medline: 29128500]
54. Kotsenas AL, Arce M, Aase L, Timimi FK, Young C, Wald JT. The strategic imperative for the use of social media in health care. J Am Coll Radiol 2018;15(1 Pt B):155-161 [FREE Full text] [doi: 10.1016/j.jacr.2017.09.027] [Medline: 29128501]
55. Baumel A. Online emotional support delivered by trained volunteers: users' satisfaction and their perception of the service compared to psychotherapy. J Ment Health 2015;24(5):313-320. [doi: 10.3109/09638237.2015.1079308] [Medline: 26485198]
56. Baumel A, Correll CU, Birnbaum M. Adaptation of a peer based online emotional support program as an adjunct to treatment for people with schizophrenia-spectrum disorders. Internet Interv 2016;4:35-42 [FREE Full text] [doi: 10.1016/j.invent.2016.03.003] [Medline: 30135788]
57. Baumel A, Schueller SM. Adjusting an available online peer support platform in a program to supplement the treatment of perinatal depression and anxiety. JMIR Ment Health 2016;3(1):e11 [FREE Full text] [doi: 10.2196/mental.5335] [Medline: 27001373]
58. Wittenberg-Lyles E, Washington K, Oliver DP, Shaunfield S, Gage LA, Mooney M, et al. "It is the 'starting over' part that is so hard": using an online group to support hospice bereavement. Pall Supp Care 2014;13(2):351-357. [doi: 10.1017/s1478951513001235]
59. Bateman DR, Brady E, Wilkerson D, Yi E, Karanam Y, Callahan CM. Comparing crowdsourcing and friendsourcing: a social media-based feasibility study to support Alzheimer disease caregivers. JMIR Res Protoc 2017;6(4):e56 [FREE Full text] [doi: 10.2196/resprot.6904] [Medline: 28396304]
60. Costello RE, Anand A, Evans MJ, Dixon WG. Associations between engagement with an online health community and changes in patient activation and health care utilization: longitudinal web-based survey. J Med Internet Res 2019;21(8):e13477 [FREE Full text] [doi: 10.2196/13477] [Medline: 31469082]
61. Patel R, Smeraldi F, Abdollahyan M, Irving J, Bessant C. Analysis of mental and physical disorders associated with COVID-19 in online health forums: a natural language processing study. BMJ Open 2021;11(11):e056601 [FREE Full text] [doi: 10.1136/bmjopen-2021-056601] [Medline: 34740937]
62. Sik D, Németh R, Katona E. Topic modelling online depression forums: beyond narratives of self-objectification and self-blaming. J Ment Health 2023;32(2):386-395. [doi: 10.1080/09638237.2021.1979493] [Medline: 34582309]
63. Oliver BJ, Forcino RC, Batalden PB. Initial development of a self-assessment approach for coproduction value creation by an international community of practice. Int J Qual Health Care 2021;33(Supplement_2):ii48-ii54 [FREE Full text] [doi: 10.1093/intqhc/mzab077] [Medline: 34849960]
64. Fjeldstad Ø, Johnson JK, Margolis PA, Seid M, Höglund P, Batalden PB. Networked health care: rethinking value creation in learning health care systems. Learn Health Syst 2020;4(2):e10212 [FREE Full text] [doi: 10.1002/lrh2.10212] [Medline: 32313837]
65. O'Donnell EA, Buccelleto AC, Van Citters AD, Nelson EC, King J, Holthoff M, et al. ConnectShareCare builds connections among care partners of people with serious illness. Alzheimers Dement 2023;19(S19):e073484. [doi: 10.1002/alz.073484]
66. O'Donnell E, Van Citters A, Buccellato A, Haines J, Cullinan A, Wilson M, et al. ConnectShareCare: a peer-to-peer online facilitated support network for care partners of people living with serious illness. Northern New England Clinical Oncology Society (NNECOS). URL: https://www.nnecos.org/22-abstracts [accessed 2024-04-10]
67. Gustafson DH, McTavish F, Hawkins R, Pingree S, Arora N, Mendenhall J, et al. Computer support for elderly women with breast cancer. JAMA 1998;280(15):1305. [doi: 10.1001/jama.280.15.1305] [Medline: 9794300]

XSL•FO
RenderX

68. Gustafson DH, DuBenske LL, Namkoong K, Hawkins R, Chih M, Atwood AK, et al. An eHealth system supporting palliative care for patients with non-small cell lung cancer: a randomized trial. Cancer 2013;119(9):1744-1751 [FREE Full text] [doi: 10.1002/cncr.27939] [Medline: 23355273]

Edited by A Kushniruk; submitted 17.10.23; peer-reviewed by N Kaufman, E Kent; comments to author 08.02.24; revised version received 20.03.24; accepted 20.03.24; published 08.05.24.

Please cite as:
O'Donnell EA, Van Citters AD, Khayal IS, Wilson MM, Gustafson D, Barnato AE, Buccellato AC, Young C, Holthoff MM, Korsunskiy E, Tomlin SC, Cullinan AM, Steinbaugh AC, Hinson JJ, Johnson KR, Williams A, Thomson RM, Haines JM, Holmes AB, Bradley AD, Nelson EC, Kirkland KB
A Web-Based Peer Support Network to Help Care Partners of People With Serious Illness: Co-Design Study
JMIR Hum Factors 2024;11:e53194
URL: https://humanfactors.jmir.org/2024/1/e53194
doi:10.2196/53194
PMID:38717809

©Elizabeth A O'Donnell, Aricca D Van Citters, Inas S Khayal, Matthew M Wilson, David Gustafson, Amber E Barnato, Andrea C Buccellato, Colleen Young, Megan M Holthoff, Eugene Korsunskiy, Stephanie C Tomlin, Amelia M Cullinan, Alexandra C Steinbaugh, Jennifer J Hinson, Kristen R Johnson, Andrew Williams, Ruth M Thomson, Janet M Haines, Anne B Holmes, Ann D Bradley, Eugene C Nelson, Kathryn B Kirkland. Originally published in JMIR Human Factors (https://humanfactors.jmir.org), 08.05.2024. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on https://humanfactors.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Gamification Approach to Provide Support About the Deferral Experience in Blood Donation: Design and Feasibility Study

Roberto Espinoza Chamorro[1], MS; Luciano H O Santos[2], PhD; Yukiko Mori[3], MD, PhD; Chang Liu[3], PhD; Goshiro Yamamoto[3], PhD; Tomohiro Kuroda[3], PhD

[1]Graduate School of Informatics, Kyoto University, Kyoto, Japan
[2]Fitting Cloud Inc., Kyoto, Japan
[3]Kyoto University Hospital, Kyoto, Japan

Corresponding Author:
Roberto Espinoza Chamorro, MS
Graduate School of Informatics
Kyoto University
36-1 Yoshida-Honmachi, Sakyo-ku
Kyoto, 606-8501
Japan
Phone: 81 075 753 7531
Email: rech91@outlook.com

Abstract

Background: Multiple studies have examined the impact of deferral on the motivation of prospective blood donors, proposing various policies and strategies to support individuals who undergo this experience. However, existing information and communications technology systems focused on blood donation have not yet integrated these ideas or provided options to assist with the deferral experience.

Objective: This study aims to propose an initial gamified design aimed at mitigating the impact of the deferral experience by addressing the drivers of awareness and knowledge, interaction and validation, and motivation. Additionally, the study explores the feasibility of implementing such a system for potential users.

Methods: We conducted a literature review focusing on the dynamics of motivation and intention related to blood donation, as well as the deferral situation and its impact on citizens. Through this review, we identified weak donor identity, lack of knowledge, and reduced motivation as key factors requiring support from appropriate interventions. These factors were then defined as our key drivers. Taking these into account, we proposed a gamification approach that incorporates concepts from the MDA framework. The aim is to stimulate the aforementioned drivers and expand the concept of contribution and identity in blood donation. For a preliminary evaluation, we designed a prototype to collect feedback on usability, usefulness, and interest regarding a potential implementation of our proposed gamification approach.

Results: Among the participants, a total of 11 citizens interacted with the app and provided feedback through our survey. They indicated that interacting with the app was relatively easy, with an average score of 4.13 out of 5 when considering the 11 tasks of interaction. The SUS results yielded a final average score of 70.91 from the participants' answers. Positive responses were received when participants were asked about liking the concept of the app (3.82), being likely to download it (3.55), and being likely to recommend it to others (3.64). Participants expressed positivity about the implementation of the design but also highlighted current shortcomings and suggested possible improvements in both functionality and usability.

Conclusions: Although deferral is a common issue in blood donation, there is a missed opportunity in existing ICT services regarding how to effectively handle such experiences. Our proposed design and implementation seem to have captured the interest of prospective users due to its perceived positive usefulness and potential. However, further confirmation is needed. Improving the design of activities that currently rely heavily on extrinsic motivation elements and integrating more social components to create an enhanced activity loop for intrinsic motivation could further increase the value of the proposed project. Future research could involve conducting a more specialized and longitudinal design evaluation with a larger sample size.

(JMIR Hum Factors 2024;11:e50086) doi:10.2196/50086

XSL•FO
RenderX

KEYWORDS

blood donation; deferral experience; Theory of Planned Behavior; Self-Determination Theory; gamification; ICT design; motivation; patient education; prototype; feasibility

Introduction

Blood Donation and Deferral Experience

Blood donation is an altruistic, socially responsible, and even a self-care activity. However, the commitment required to participate can be deterred by uncomfortable experiences, negatively affecting motivation to donate [1]. Preventing such experiences and reducing their impact could preserve the intention and willingness of citizens to continue donating blood.

One significant deterrent is the deferral experience [2], which means being disqualified from donating blood based on eligibility criteria. It is especially impactful in young and first-time donors, often resulting in abandonment [3]. Negative emotions are commonly reported during the deferral process [4]. Such negative experiences can deter potential donors when shared with their social circles [5]. Similarly, studies consistently identify negative interactions with staff, feelings of rejection, and confusion about deferral reasons as primary factors reducing return intention and motivation [6]. Some deferred donors misunderstand their deferral conditions, erroneously believing they cannot donate for longer periods or even permanently [7]. Communication and information gaps contribute to these misconceptions [8].

Challenges and Potential Solutions

To counter deferral effects, strategies such as enhanced communication, clear deferral information, and targeted recruitment show promise [4]. However, these solutions require substantial planning and resources, often unavailable to many blood centers. In that regard, information and communications technology (ICT) platforms, inspired by successful implementations in health promotion and telemedicine [9,10], could facilitate such support. As existing apps focus mostly on the donation itself and on supporting citizens for their next donation [11,12], there is an opportunity to offer unique value by tailoring systems to address the deferral experience. However, to that end, understanding the psychological responses and specific needs of deferred donors is crucial. Temporary deferrals necessitate motivation, health improvement, and eligibility [13], whereas permanently deferred donors (those unable to donate anymore) could still contribute indirectly through activities such as promoting donation.

Regarding the motivation topic, an approach that has gained notoriety is gamification, which involves the use of game-design elements in nongaming contexts [14]. Some blood donation centers worldwide use gamification, rewarding donors with badges, gifts, and certificates [15]. Furthermore, government blood donation apps in countries such as the United States and Canada have integrated gamified elements into their ICT services, aiming to boost donor motivation [16,17]. Although the impact on blood donation is yet to be studied, gamification has proven effective in therapy commitment and health self-monitoring [18,19]. Considering this, gamification holds promise for increasing motivation in the blood donation context.

However, motivation is not the only factor to consider for a possible proposal. Previous studies have identified various deterrents, stemming from deferrals, health conditions, or environmental factors, which influence citizens' future intentions and behaviors [4]. In this study, we reviewed previous findings, as well as the results of a preliminary survey, to identify pertinent topics and form the foundation of our proposed design for an ICT system that aims to support (also) deferred donors.

Theory of Planned Behavior and Extensions in Blood Donation

The Theory of Planned Behavior (TPB) [20], an extension of the Theory of Reasoned Action, asserts that specific behavior is determined by intention, influenced by attitude, subjective norm, and perceived behavioral control. In the context of blood donation, the TPB proposed that positive evaluation of the act (attitude), social expectations (subjective norm), and belief in individual control over donation (perceived behavioral control) dictate the decision to donate.

Previous studies have found that the TPB explains between 32% and 50% of the variance in intention and 27% and 36% of the variance in behavior [21,22]. To enhance predictive power, the framework was extended due to the inconsistent link of the subjective norm [23]. In blood donation, additional constructs were incorporated based on psychological differences among nondonors, novices, and repeat donors. Moral norms, descriptive norms, past behavior, and self-identity were included as predictors [20,24,25].

Systematic reviews have shown that self-efficacy, donor identity, and anticipated regret have medium positive effects on both intention and behavior. Conversely, deferral has a medium negative impact, leading to a decrease in subsequent donations among experienced donors [26,27]. Past behavior or habit explains 19% additional variance in blood donation behavior for those donating 5 times or more [28]. Habit, suggested to be context bound, is viewed as an external motivator, whereas self-identity, which pertains to one's role in society as a blood donor, is defined as an internal motivator [29-31]. Both habit and self-identity significantly influence repeat blood donation behavior, with past behavior likely forming identity [23].

Self-Determination Theory and Motivation and Gamification

As the TPB applies mostly to situational-level intentions [32], blood donation studies primarily rooted in the TPB have expanded their scope to incorporate Self-Determination Theory (SDT) in the last decade [33,34]. SDT, a theory of human behavior and personality development, emphasizes social-contextual factors supporting human growth through satisfying basic psychological needs for *competence* (effectiveness of my actions in my current environment), *relatedness* (social involvement and relation with others), and

XSL•FO
RenderX

autonomy (internal need to be responsible for your own meaningful choices) [35]. It proposes that internally motivated behaviors persist, while external motivations can become internalized under appropriate socioenvironmental conditions [36].

SDT categorizes behavior on a continuum from amotivation (nonregulated behavior) to extrinsic motivation (external to integrated regulation) to intrinsic motivation (intrinsic regulation). Extrinsic motivation refers to acting in a certain way or doing a specific action because it leads to a separable outcome or reward. By contrast, intrinsic motivation refers to acting in a certain way or doing a specific action because the act itself is inherently satisfying. Integrating the TPB and SDT, studies have revealed that SDT's motivational orientations explained an additional 14% of the variance in blood donation intention compared with TPB-only models [37]. Amotivation had a negative direct effect on intention, while external motivation had no overall effect on intention but a positive effect on amotivation [38]. By contrast, introjected regulation had positive direct and indirect effects on intention, and autonomous motivation predicted intention directly and via attitudes, subjective norms, and perceived behavioral control [33,38].

Gamification Concepts and Frameworks

As SDT discusses the impact of motivation on behavior, it was considered the foundation for implementing gamification, as it does not aim to directly affect an outcome, but to change a target behavior (by affecting psychological factors) that can lead to that outcome [39,40]. To achieve this, the system can utilize its various design components, as outlined by the Mechanics, Dynamics, and Aesthetics Framework (MDA) [41], which served as the primary reference for our study. The framework comprises mechanics, which encompasses specific game components such as data representation and algorithms; dynamics, which refers to the interactions between these mechanics and player inputs over time; and aesthetics, which aims to elicit desirable emotional responses from players when they engage with the game system. These components are integrated to drive either extrinsic or intrinsic motivation, considering the targeted changes in human behavior [42-44].

Extrinsic motivation can drive behavior but may fade without external rewards, while intrinsic motivation leads to long-term positive effects on intention and behavior [45,46]. Thus, most gamified approaches recommend prioritizing intrinsic motivation in the design process. In that regard, users can be categorized according to the recognized characteristics and that drives them in the gamified implementations [47]: socializers (motivated by *relatedness*), free spirits (motivated by *autonomy*), achievers (motivated by *competence*), philanthropists (motivated by *purpose and meaning*), players (motivated by *rewards*), and disruptors (motivated by *change*). The players and disruptors categories can be further divided according to their behavior.

Considering the previous concepts and relationships of gamification and SDT, DiTommaso and Taylor [39] defined a framework in which they propose the following steps for design: discover the reason to gamify, identify players' profiles and motivational drivers, set up goals and objectives, describe skills and desired outcomes, and playtest among others. Another design framework with similar foundations is the Six Steps to Gamification [48], which also takes influence from the MDA. It proposes the following steps: definition of business objectives, target expected behavior, description of players, design of activity loops, do not forget the fun, and deploy appropriate tools. Although not domain specific, these adaptable frameworks can guide gamification projects and were also used for reference in our study.

Deferral Experience and Effects in Return Rate

From the literature review, we chose to focus on recurrent and impactful issues related to the deferral experience, especially the ones that aligned with the constructs from the TPB and SDT. For example, the construct of self-identity (blood donor identity in this case) from the extended TPB can be associated with the negative feelings from a deferral. More specifically, a deferral, which can generate a feeling of rejection in the unsuccessful participant [13], can threaten the citizen's self-perception as a capable blood donor (identity), as the inability to participate diminishes their possibility of building experiences and forming a habit (especially in the cases of new and young donors). Similarly, confusion and misunderstandings in deferral make a successful blood donation seem more complex and difficult than it is, affecting citizens' perceived behavioral control (TPB construct). As indicated by Gemelli et al [1] and Hillgrove et al [13] negative experiences can reduce the motivation for future involvement, particularly for long-term or permanently deferred donors, eroding their sense of self-efficacy.

To further explore the relationship between the deferral experience and intention, we also took into account the findings of a preliminary survey involving Japanese citizens [49], in which a total of 208 participants were recruited. In the survey, the dependent variable was "Intention to donate again after deferral" (a 6-point Likert scale question with the values 1=not anymore, 2=not for a while, 3=I don't know, 4=maybe, after a while, 5=yes, unless rejected again, and 6=yes, I would). Citizens were asked whether they heard or knew about the concept of deferral and whether they had experienced a deferral case, as well as their future intention in a possible deferral scenario. The results implied a possible relation between deferral and reduced intention to donate (following previous studies). However, the data also suggested a positive relation between preventive awareness of the deferral experience and intention to donate. Donors and nondonors who had knowledge about the deferral concept indicated higher intention of future participation even after a possible deferral scenario.

Objectives

Considering the literature review, we focused on recurrent issues that could be addressed with a gamification approach, taking into account the connections between the deferral experience, their issues, and motivation. The topics we chose were as follows:

- Lack of knowledge about deferral: Some of the negative feelings appear because citizens are not knowledgeable of the topic, are not retaining the information, or have misunderstood it.

XSL•FO
RenderX

- Weak donor identity: citizens feeling rejected and lacking validation.
- Reduced motivation: citizens losing interest in addressing the deferral reason or losing interest in contributing to the future.

Additionally, considering the evaluated strategies to mitigate the negative impact of deferral from the analyzed literature [4,50], we defined the main drivers for our approach. First, to provide awareness and knowledge about deferral by making learning interesting to the citizens (*awareness and knowledge*). Next, to increase the scenarios of interaction and validation for deferred donors to nurture their identity (*interaction and validation*). Lastly, to provide motivational drivers for deferred donors to regularly engage in activities related to blood donation (*motivation*).

After that, we worked on the design of activities that could be implemented with the gamification framework while targeting the drivers selected regarding the deferral experience. After completing the initial design, we implemented a prototype with basic features and integration for a feasibility study, collecting feedback about the usability and receptivity of potential users to discuss the future value of the idea of offering a service regarding the deferral experience, our proposed design, and its implementation.

For this study, we explored the following research questions (RQs):

- RQ1: Will our gamified design that focuses on the previously mentioned drivers with regard to the deferral experience in blood donation have a positive reception from potential users?
- RQ2: Will our initial prototype implementation of the design be considered usable and useful in its current iteration?

Methods

Conceptualizing a Gamification Approach for the Deferral Experience

Overview

In this study, we are adopting an approach similar to the gamification frameworks mentioned previously [40,41,48], while also taking into account the unique requirements of individuals in blood donation. We have adapted the steps and elements of these frameworks to provide support specifically addressing the deferral experience and focusing on the main drivers mentioned.

Definition of Approach Objectives

We redefined our target users to include not only deferred donors but also regular donors and potential donors who might face deferral in the future. Our focus broadened to cater to anyone interested in the topic, aligning with our objective of providing deferral support. We concentrated on 3 main issues: lack of knowledge about deferral, weak donor identity, and reduced motivation, translating these into drivers for our gamification approach: *awarenessand knowledge*, *interaction and validation*, and *motivation*.

Target Expected Behavior

The next step was to define the citizens' expected behavior when interacting with our proposed gamification implementation. For our approach, we wanted the design to nurture the drivers, and as a consequence, possibly affect future intentions.

For *awareness and knowledge*, we expected users to engage in educational activities that both teach them about and test their understanding of the deferral experience and strategies for improvement. For *interaction and validation*, we expected users to get involved in discussions, in sharing experiences, and in supporting one another, improving the sense of community. In terms of *motivation*, our goal was to encourage users to access the system regularly, ideally once or twice per week, considering the prolonged pace between blood donations.

Description of Users

For our target group, while we initially expected to focus on the deferred donors, the results from the preliminary survey guided us to design the service as a preemptive one (including regular donors and nondonors), to nurture the identity of the users and prepare them against a deferral scenario. Designed primarily for young citizens (20-30 years) yet accessible to older individuals, the approach incorporated specific design elements reflecting the regional context (Japan). However, the core of the approach was intended to be adaptable, considering possible future adaptations for other regions.

In the context of the gamification approach, considering that the potential users (citizens) would not have the same goals or motivations (following the connection with SDT), for this study, we focused on targeting the players, the socializers, the free spirits, the achievers, and the philanthropists.

Design of Activity Loops

Macrolevel Progression Loops

The gamification approach aims to motivate citizens, particularly deferred blood donors, to stay engaged with blood donation–related activities. Although encouraging future donations is the ultimate goal, maintaining interest in the topic and promoting contributions to other related areas are also crucial. The design focuses on creating macrolevel progression loops for the drivers of *awareness and knowledge*, as well as *interaction and validation.*

Initial Outline

User progression is represented through levels. Levels increased based on experience points earned from various activities. Points earned could be exchanged for basic title characters. Special characters are unlocked as users progress, with higher levels requiring more points for unlocking. Higher user levels unlock additional activity options, which yield more points.

For the microlevel, we first defined some basic loops for the foundation of the design. For example, one of the initial hurdles considered was that, independent of any learning or social activity that could be designed, their value would not be

XSL•FO
RenderX

achieved if the users were not motivated to access the ICT system. In that regard, we considered a simple loop of providing a reward to initially push the user to use the system: if the user logs-in to the system, they receive a message about their current streak and earn some points. Users will earn more points according to how often they connect to the app and how high is their level. With regard to *awareness and knowledge*, we aimed to make the users both learn about deferral and review their current knowledge. For this purpose, the initial idea for this loop was that as users learn more, they can face harder challenges. And the more successful they are, the more complex information they will be taught. With regard to *interaction and validation*, we aimed to provide some activities in which users could interact with other users, and the more interactions and levels the user has, the more options of interaction would be available to them.

However, we needed to solidify the ideas for the microlevel. To achieve this, we opted to elaborate on the design with greater detail. We chose to do this by following the MDA framework, first from the user perspective, then transitioning to the designer perspective to finalize the activities' design.

Definition of MDA Aesthetics

For the driver *awareness and knowledge*, we aimed to nurture a habit in the users of learning about deferral. For *interaction and validation*, the expected behavior was to generate regular engagement in the users. To that end, specifically for the players, we first selected *submission*, which means the design would allow users to interact with the system as a pastime. Our goal was to present a variety of activities offering rewards and collectible items to enhance user enjoyment. However, this approach may heavily rely on extrinsic motivations, potentially overshadowing the altruistic aspect of blood donation. Thus, we needed to be cautious in its implementation to avoid solely focusing on rewards. To address this, we selected *fellowship* as a social framework to appeal to users who value social experiences.

We considered possible ways to make users interact with others, possibly in cooperation or competition. *Challenge* (experience as obstacle course), *discovery* (experience as uncharted territory), and *expression* (experience as self-discovery) were also chosen as they are more related to intrinsic motivations, which we wanted to favor over the extrinsic motivation, which was aimed to be used only as the trigger for the conduct of the users.

Definition of MDA Dynamics

We initially drafted dynamics outlines to connect the drivers and the aesthetics. For instance, in terms of *awareness and knowledge*, users could opt to heighten the difficulty of their learning process, introducing an element of risk that could generate a *challenge.* Additionally, we explored the possibility of randomizing the information users received, with variations based on their actions within the environment, thus fostering a sense of *discovery.*

By contrast, for *interaction and validation*, we aimed for users to be able to choose the type of recognition they would get, allowing for *expression*. They should also decide what they could share with others and try to encourage them to perform certain actions, creating *fellowship*. From these initial ideas, we expanded into more detailed dynamics in the designer perspective iteration of the MDA.

Definition of MDA Mechanics

Generic mechanics are introduced, incorporating points, levels, and characters for onboarding. Points served as rewards for participating in different activities (the amount was adjusted per result), to create a sense of progression (the historical record was tracked to calculate the current level of the user), and to be used as a currency in the system. Levels were also used for progression. They increased according to the number of participations, providing recognition and incremental rewards. They were used as a certain multiplier in the activity rewards and to unlock new and special characters in the exchange store.

Characters were chosen as part of the representation and recognition of the users, being the main extrinsic reward of the gamification approach. However, they were integrated to appeal to both extrinsically and intrinsically motivated players, aiming to reduce the dependency on the extrinsic component. For example, with customization, they would target free spirits; if they were collectible, they would target players and achievers. As general rules, every registered user was provided with the same starting character; they could acquire more in the shop by exchanging the points they collected through the activities; they could also upgrade (defined as "evolve" inside the app) them by exchanging multiples of the same one. One character at a time could be selected to use it as their icon in their social activities, and characters would change their appearance if the user stood inactive in the system for more than a week.

Some social interaction components were included, such as a comment section and a simple feed wall for users' posts. Both of them had an upvote or downvote mechanism for users to indicate their relevance or popularity. A certain degree of user anonymity was incorporated to reduce possible social burdens of participants when creating content. However, for regular comments, the app showed their current character (and title) and their username. The main posts were put on hold until approved by an administrator, to reduce possibly harmful or misleading content; however, regular comments did not have this restriction. These mechanics aimed to engage the socializer, the free spirit, and the philanthropist types of users.

After this first iteration, we started with the designer perspective, in which we focused on linking all the previous concepts together, defining the more specific activities available in the system for the users.

MDA-Based Features and Feedback Loops

Finally, we describe the design of our proposed features for the gamification approach, integrating all the previous considerations and concepts.

The first feature we defined was the "Login Reward." Usage of blood donation apps tends to be low because of the timed nature of donating blood. However, to handle learning and engagement, as part of the onboarding, we chose to encourage users to interact with the system more often. To that end, we rewarded points if users log-in to the app regularly with up to

XSL•FO
RenderX

5 rewards per week, increasing the amount per consecutive access. We linked the reward to the level mechanic, providing additional points according to the level. Regarding the *awareness and knowledge* driver, we also included a message of advice and information regarding blood donation deferral. As dynamics, users could choose to access the app as usual or connect more times to increase their multiplier. Besides, as the level was linked to the rewards, users could choose to increase their level through other activities to receive more points. However, as users were not forced to read the advice message, we connected it with other activities to create the intended aesthetics and more complex activity loops.

The next features we defined, *quizzes* and *social poll*, were aimed to be connected with the broader activity loop and the *awareness and knowledge* driver. Quizzes have been implemented in other blood donation apps, so we included additional mechanics to make it less extrinsic, create new dynamics, and reach the intended aesthetics. We incorporated a life mechanic that resets daily, along with a difficulty level that becomes unlockable as users progress through levels, giving less incentive to guess the answer while also providing a higher risk-higher reward choice to more expert users. Additionally, feedback was provided according to the result, either congratulating the user for their right answer or guiding the user on the mistake. Furthermore, we connected the questions with the content shared through other features, so invested users will feel rewarded for learning on their own. Similar ideas were considered for the content of the *social poll*, but some mechanics that could allow for social interaction were included. Once per week, users could vote between different facts related to blood donation deferral, according to what they felt was the most interesting one. At the end of the week, users were notified of the most popular choice, and the ones who chose it were able to claim reward points. If desired, users could either discuss outside or through the app to try to get information about other users' preferences or to coordinate a specific choice for benefit. Additionally, previous results and facts were accessible, so users could review the content and discuss it for self-learning or connection with users.

The *news sharing* feature was also connected to the previous features and the *awareness and learning* driver, as content shared on the former would be used in the latter ones. Users could like their favorite entries and could comment about them. Comments or replies from administrative users had a special identifier while regular users had the default. Administrators would try to reply to important questions, but the content of the comments or discussions was up to the users, giving them freedom for communication.

Similarly, to provide more options for the *interaction and engagement* driver, we defined the *posts* feature. Users could create posts for discussion (questions, anecdotes, suggestions, among others). If approved, the posts were shown in the app anonymously, displaying their current relevancy score. Every week, users who created new posts with high relevance would be rewarded points. While posts were defined to require approval by the administrator users, that restriction was not included in the evaluation. Posts would show in the user's feed, by order or relevancy and created date. We aimed to reward users for meaningful content, which in itself could motivate the participation of other users in the discussions. Besides, as the more relevant ones would be highlighted in the app feed, it could provide a sense of self-worth by knowing that one's content received a good reception from the community or that it provided value to the community, eventually motivating them to participate again in the discussions.

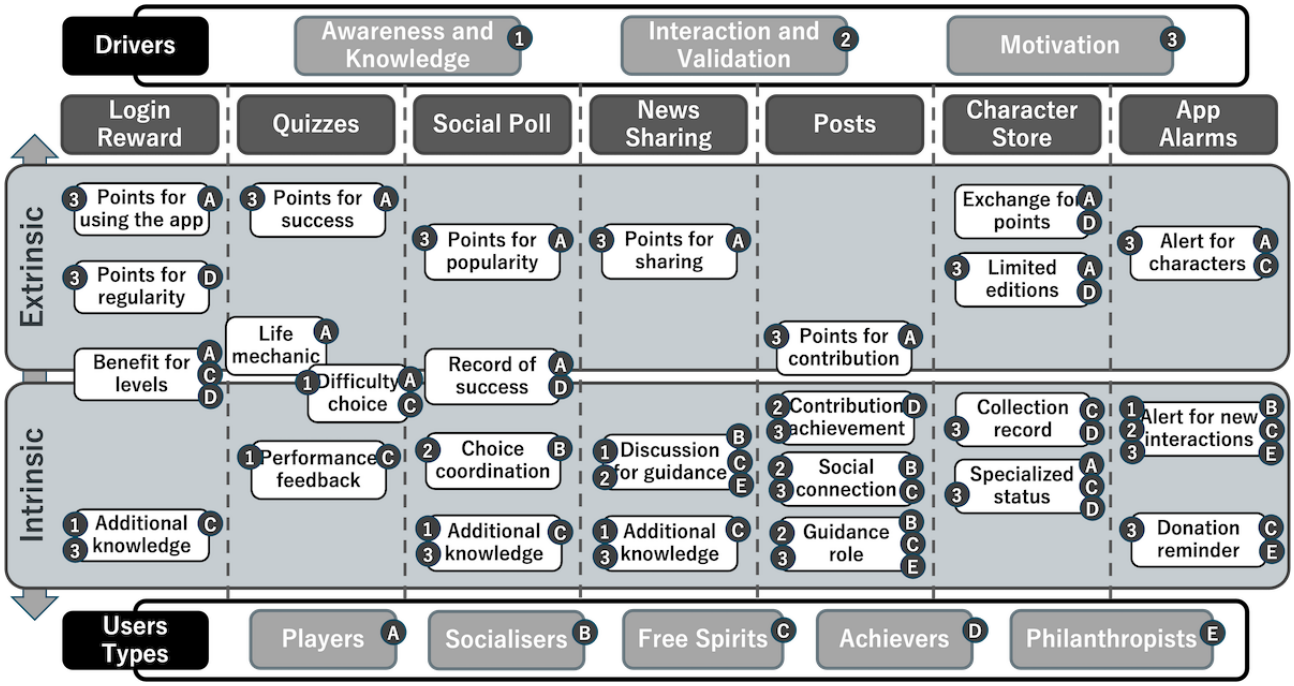
The next feature was "Application Alarms," aiming to provide users with some mechanics that could support their preferences. Users had the option to enable up to 3 types of notifications: notifications when new characters were implemented, notifications for news and discussions, and a reminder of the calculated end of the deferral period. The aim was for users to voluntarily choose to get informed about their topics of interest within the app.

Finally, for this initial scope of the design, we included features that, while not creating a proper loop by themselves, were required to connect the previous features and their loops. The first one is the "Character Store," in which users can exchange their points for available characters. The list of characters was updated in a regular schedule, with new characters being highlighted, while locked characters had a gray background. The store showed the required level, price, and current amount collected for each character. Characters being collectible were used as an extrinsic way to motivate the users to keep getting points through the other activities. By contrast, the upgrade option was linked to the title achieved by the user, which meant a special title for their effort. Users had the option to concentrate their points on either one objective or the other or to participate as much as possible to pursue both simultaneously. The other feature was the "Profile," in which users had access to their stats (level/points), their character collection (including the upgrades and selection), and additional settings for the account.

Some of the mentioned intended connections between the drivers, the users, the features, and their gamification elements can be seen in Figure 1, which provides a more general outline of what we aimed to integrate as part of the activity loops.

XSL•FO
RenderX

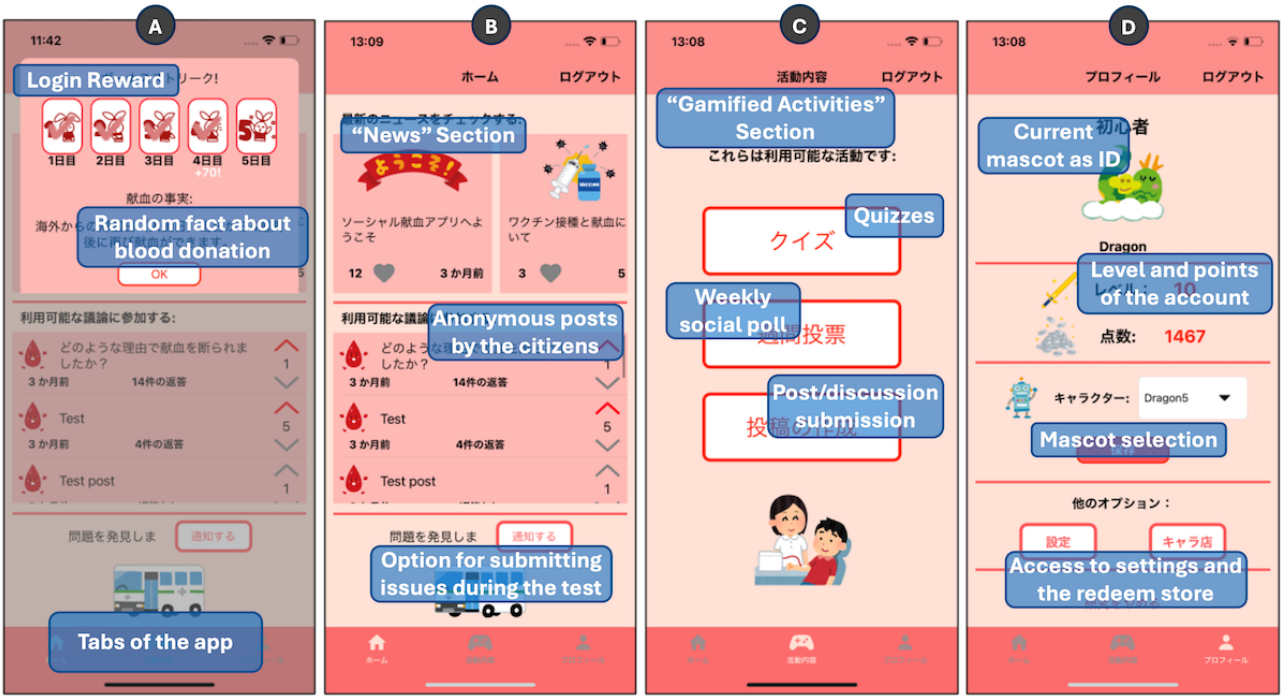
Figure 1. Outline of connection between drivers, users, features, and gamification elements.



Deploy of the Concept

The proposed design was implemented in a basic mobile prototype, named as "Social Blood" app, to encapsulate the idea of a more interactive role from the citizens in blood donation. For the icon and the other illustrations of the prototype, public domain images were selected from the Japanese web page Irasutoya [51] for the test deployment. The main screens of the app are displayed in Figure 2.

Figure 2. Screenshots of the main sections (Japanese version) of the prototype: (A) Home tab with the "Login Reward" as a pop-up; (B) Home tab screen; (C) Activity tab screen; and (D) Profile tab screen.



A welcome screen was created for user registration with either an email or Facebook account. An additional functionality (In-app Survey) not related to the design was included for data collection. The app would check for surveys requested by the researcher or the staff and ask the user to answer them. Once the pending surveys were completed, the user was redirected to the main part of the app. If the user was accessing the app for the first time in the day, the "Login Reward" feature was shown to them.

A "Home" tab was created as the main interface available to the user. This section included features related to both learning and interaction, such as "News" and "Posts." The user could

XSL•FO
RenderX

check the number of likes and votes of any entry on this screen. An additional option not related to the design, "Error Report," was also considered in this screen, to let users notify the administrators if any issue was found during the use of the app. An "Activity" tab was created to include the features that support learning. Users can access the "Quiz" and "Social Poll" features on this screen. Users could create a discussion post on this screen. Finally, a "Profile" tab was also created to show to the user information regarding the gamified elements of the app. This section of the app connects to the "Character Store" feature and to the "Settings Screen" screen, which includes the "Application Alarms" feature.

Recruitment of Participants

To collect initial feedback regarding the prototype for its usability and acceptance by prospective users, a survey was performed with volunteers recruited on social networks. A digital flyer was posted with details of contact for the interested parties (Multimedia Appendix 1). Prospective participants were required to have an iPhone (Apple Inc.), be between 20 and 50 years old, and live in Japan for at least the last two months. Participants were recruited from June 17 to July 2, 2021. No incentives were used for the recruitment. Interested citizens received a Google Form (Google LLC/Alphabet Inc.) with the informed consent details and registration (Multimedia Appendix 2). If they signed up for participation, they later received an email with the following: a link to download the app, the user manual of the app (Multimedia Appendix 3), a list of main tasks to complete inside the app (Multimedia Appendix 4), and another Google Form link that contained an anonymous survey (Multimedia Appendix 5). Participants were asked to first download the TestFlight app from the Apple Store, and from there, install and use the approved version of the research prototype for a few days. They could then follow the tasks and complete the anonymous survey either through the app or through the Google Form once they deemed their test as completed. They could test the app and submit their answers to the survey until July 11, 2021.

Ethical Considerations

The study focused on collecting preliminary feedback (usability and acceptance), so no sensible information was stored, and no risk nor effect was involved for the participants. With those points in consideration, considering the guidelines of the Kyoto University Graduate School and Faculty of Medicine Ethics Committee, it was not required to apply for ethical approval.

Evaluation Details and Data Collected

Participants were instructed to attempt to complete the list of primary tasks outlined in the prototype app (Multimedia Appendix 4). Hereafter, these tasks are referred to as follows: Register and log-in (1. Log-in), fill out the survey (2. Survey), interact with news posts (3. News), interact in the discussion posts (4. Discussion), participate in the quiz activity (5. Quiz), participate in the weekly poll activity (6. Poll), submit a simple post (7. Post), acquire a new character (8. Buy), upgrade a new character in the Character Store (9. Evolve), select a new character for your profile (10. Select), and finalize their session (11. Log out).

In the Google Form, participants were asked to answer the following sections: demographic questions (age and gender), difficulty of task completion (questions about the previously mentioned list of tasks), System Usability Scale (SUS), and follow-up questions divided into acceptance questions (a Likert scale of 5 items) and opinion questions (free-text answers). Further details of these questions are provided in Multimedia Appendix 5. Additionally, participants were given a contact email to ask for support in case they had issues during the testing.

Data Analysis

Statistics of mean, SD (σ), and standard error of the mean (σ_M) were calculated for the average of task difficulty and the final SUS results. The SUS score per participant was calculated according to the standard assignment of points per type of question, which included positive and negative values [52,53].

Qualitative answers were grouped and summarized (if possible) following a simple semantic approach: we grouped the answers for each question and summarized the main ideas according to positive or negative feedback regarding the topic of the question. For the questions regarding the status of the app, we used the labels interface, functionality, and gamification to group the answers. The same categorization was followed for the questions regarding suggestions and improvements. Additional comments were not segregated but were individually considered and described, provided they were not redundant.

The analysis of the questions was carried out after the submission deadline for participant results had elapsed. Only submissions that were completed and received before the deadline were taken into account for the final analysis.

Results

Overview

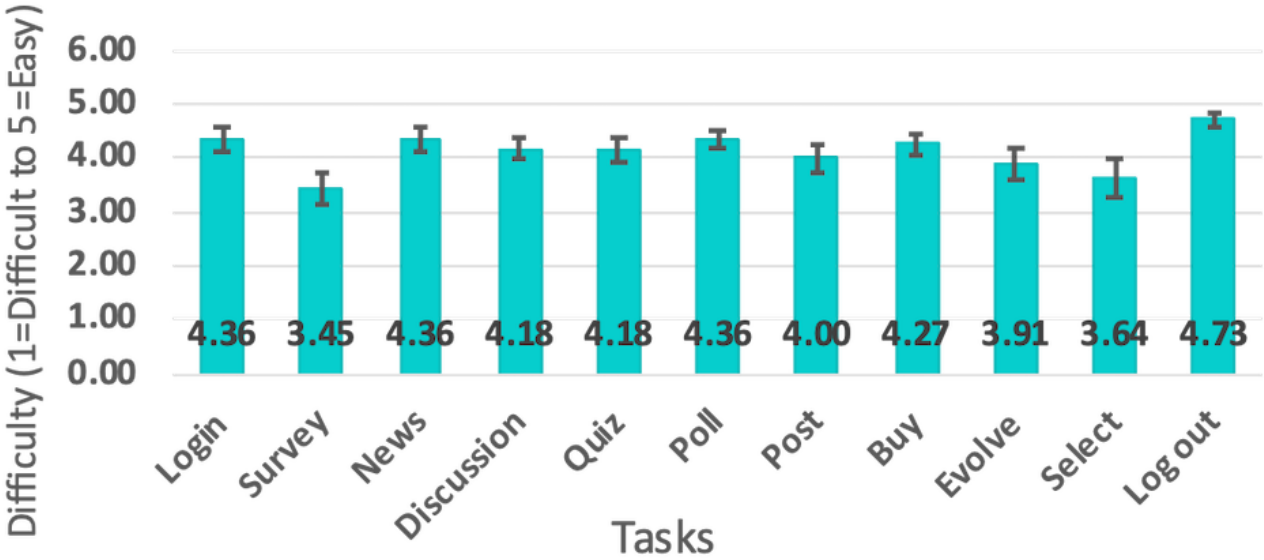
From May 17 to July 2, 15 participants were recruited for the preliminary evaluation. A total of 13 participants created an account for the app and 11 participants submitted the final form. The full content of the answers is available in Multimedia Appendix 6. Participants ranged in age from 20 to 34 years. There were 10 male and 1 female participants. Descriptive statistics were used regarding the demographic variables of the participants.

Difficulty of Task Completion

Only 4 participants asked for support during the period of the evaluation. Questions were related to tasks 2, 7, 8, 9, and 10. Only tasks 9, 10, and 11 had 1 case each of not being completed. Regarding the difficulty level of each task, the results indicated that all of them were relatively easy to complete (4.13 average). The tasks that were considered the most difficult were task 2 (Completing the in-app surveys) and task 10 (Selecting a new character), as shown in Figure 3.

XSL•FO
RenderX

Figure 3. Subjective results of difficulty of each task.



System Usability Score

The SUS evaluation of our proposed app, as seen in Multimedia Appendix 7, showed a final score of 70.91 (scale 0-100, with 100 being the best usability), slightly above the average SUS score of 68 (C grade, percentile range of 41-59). The highest SUS score received by participants was 95, while the lowest score was 30.

Regarding the score per question, item 1 (“I think that I would like to use this application frequently”) showed the average lowest score from all the lists, with a value of 3.18. The highest score was for items 3 (“I thought the application was easy to use”) and 7 (“I would imagine that most people would learn to use this application very quickly”), with a value of 4.00.

Follow-Up Questions

Acceptance and Qualitative Questions

Regarding the acceptance questions, participants responded positively to the app, expressing interest in its concept (3.82), likelihood to download it (3.55), and likelihood to recommend it to others (3.64).

For the qualitative questions, we summarized the answers for the main topics of the survey.

Goals of the App

Most participants considered a blood donation app concept useful or helpful. From them, 2 participants highlighted the possible impact of the deferral experience. The other 2 participants focused on the service being an app as a core value of the project.

Factors That Could Motivate Usage

Three participants emphasized that being aware of how they can contribute can help maintain their motivation; 3 participants mentioned interaction with others and popularity of the app as their motivation; 3 participants focused on the gamification aspects as one of their factors; 2 participants highlighted the social components as their drivers; and 2 participants indicated possible personal benefits for motivation.

Preferences About the App

Three participants liked the interactive possibilities of the app; 2 participants indicated the Quiz as their preferred feature; 3 indicated sharing and discussing as their favorite activities; 2 mentioned liking the activities involving characters; 1 participant indicated to like the interaction in general; and 1 participant indicated that they liked the aesthetic of the app the most.

Weaknesses of the App

Some participants recommended support of more languages so more citizens could benefit from the app. One participant indicated that the Quiz activity required improvement but did not specify reasons. Two participants indicated that the character functionality could be improved. One participant complained about the compulsory survey in the app because of its duration. One participant felt that not all the gamification features were connecting well with the goal of the project.

Current Status of the App

Only 1 participant mentioned that the current features might not be sufficient to support the goals of the app. They mentioned that while the app can be used to support deferred donors, it might not motivate them to promote blood donation. The other participants provided feedback regarding adjustments or fixes for the current version of the app (Textbox 1).

XSL•FO
RenderX

Textbox 1. Feedback regarding adjustments or fixes for the current version of the app.

1. Regarding the User Interface

   Possibly change the color palette of the app or allow theme selection, as the red color might create discomfort. Implement support for the "Dark" mode, as it created issues with the color of text in the News and Post features.

2. Regarding Functionality

   Reduce the length of the In-App Survey. If possible, implement support for more languages, as it could help international students who want to donate.

3. Regarding Gamification

   Adjust the point requirement for characters, as it was too high in the test. Move the character selection option to a grid, so users can look at their whole collection when choosing. Add feedback messages in the Quiz about the points acquired.

Improvements and Suggestions

Participants were also asked about what they wanted to see for implementation in the future (Textbox 2).

Textbox 2. Participants' expectations for implementation.

1. Regarding Functionality

   Consider the inclusion of a feature to find locations for blood donations. Consider the inclusion of features to share the news and discussions on social networks. Allow to link or upload videos in the comments.

2. Regarding Gamification

   Consider adding the creation of groups or friend requests. Consider adding a "Gacha" option to acquire exclusive characters. Consider adding a ranking or certificate, similar to what is implemented in "Duolingo" [54].

Additional Comments

Some concerns about the information allowed in the Post feature were mentioned, as it could be nonrelated or harmful to the users. The usefulness of the app would be higher if medical institutions could provide information within it. It was suggested to highlight to the users the core goal of the app during the registration. It was also suggested to allow donors to know when their blood is used, as it could help to motivate them to continue to donate blood.

Discussion

Principal Findings

Current ICT services in blood donation aim to improve the citizens' experience but do not focus on the deferral experience and its effects on prospective donors. This paper contributes to the field by debating the viability of implementing a system focusing on deferral and proposing a novel design to expand the concept of contribution and identity in blood donation. Our study indicates a missed opportunity in current services related to deferral. Potential users seem interested in an app supporting them in this area, and social gamification could make the role of a blood donor more approachable. However, our results, although slightly positive, require further validation due to the limitations, leaving room for discussion regarding the gamified design and the implemented prototype.

Proposed Gamified Approach Reception and Shortcomings

- RQ1. Will our gamified design that focuses on the previously mentioned drivers with regard to the deferral experience in blood donation have a positive reception from potential users?

Participants' favorable responses (average Likert scale score of 3.82) and positive opinions about the proposed functionality gave us an initial indication that the proposed project could be beneficial for the community. These results seem to align with ideas and concepts previously discussed in other studies. Previous studies discussed the relationship between knowledge of blood donation and intention to donate blood. However, only a couple of reviewed studies evaluated the ratio of knowledge regarding deferral. From our preliminary survey in Japan [49], 33% of nondonors did not know about the concept of deferral, with an additional 11% also unaware of the concept. Similar results were shown in [55], in which 90% of the participants never heard about the "donor deferral" term. This unawareness regarding deferral could be related to the positive response from the participants in our project, as either it introduced them to a new but relevant concept or it emerged as a service that could be valuable because of the low level of current support, which can be considered from their answers in the open questions. Furthermore, as participants expressed their positive intention to download the app (3.55) and to recommend it to others (3.64), the results suggest that there could not only be an interest but also an emerged necessity that has not appeared before because of the lack of awareness.

However, the current data are insufficient to reach a proper conclusion about the project acceptance, not only because of the small sample but also because of the scope of the participants, as it is not a proper representation of the target population. Additionally, the positive reception from the users could have been influenced by the Hawthorne effect [56], as

XSL•FO
RenderX

the participants were aware of being part of an experiment, and the topic was related to a social contribution project. In this regard, a higher-scale study is required for further validation and analysis to decrease the effects of noise in the data and allow for more significant results.

Regarding the gamification aspect and its value, while the mentioned results were positive, their approval could have been related more to the goal of providing support. We delved into the comments of the participants about the design itself for possible conclusions. When asked about motivation to use the app and its best feature, some participants did indicate that the gamified aspects caught their interest and could even be driving motivators, highlighting the characters as part of it. These answers seem to suggest that the gamified components can play a role in, at least, capturing the interest of potential users. However, more detailed data are required to determine how beneficial is the integration of the gamified concepts in our proposed project. For example, asking participants for specific reasons why a gamification feature seems motivating to them or why it might feel discouraging. Besides, an additional evaluation regarding the impact is considered, as the value of the proposal can be confirmed if a positive effect can be determined. Comparing which features have more or less effect could also be important, as it could allow for the identification of factors to consider for future ICT-gamified implementations in blood donation.

Currently, while we discussed the importance of integrating the type of users, their motivations, and the MDA elements to nurture the drivers of interest, we have no specific data to indicate if our design has the desired effect or not. The data limitation becomes important with our goal of nurturing the intrinsic motivation of the citizens (prospective users), as we cannot recognize if the potential interest is related to the components that nurture the intrinsic motivation or the ones that do so for the extrinsic motivation. Analyzing some of the comments, most of the positive focus was on the *quiz* and the *characters*. These features, although designed considering an activity loop that could nurture intrinsic motivation, might not reach that goal in their current state. This weakness appears to be echoed in the feedback from 1 participant, who expressed dissatisfaction with the current state of the app, feeling that it falls short of achieving our design goals and lacks sufficient integration of features. As some participants showed interest in the social activities of the design (which are more related to intrinsic motivation), it might be worth it to redesign the current gamified activities to incorporate and integrate social components as part of the progress of the users.

We previously mentioned that some restrictions should be considered in a system related to blood donation, as some interactions could clash with the altruistic nature behind the donation act. To address that complexity, having a deeper understanding of game design itself is required. Learning from different and successful implementations of player interactions in game environments can lead us to a design that can properly nurture prospective blood donors' social motivation. From the case studies in *The Gamification of Learning and Instruction Fieldbook* [57], an interesting idea is the implementation of specific types of leaderboards that encourage various forms of participation, thereby creating a stronger activity loop. Building on that concept, although we aim to steer clear of incentivizing competition in donation participation, we could adapt similar interactive mechanisms to enhance engagement in learning activities. For instance, in the Quiz activity, introducing a monthly leaderboard alongside corresponding achievements could offer users more personalized motivation compared with simply rewarding points. Another intriguing option could involve allowing users to accumulate questions they have answered correctly, which they could then use in a soft-competition interaction. In this scenario, users could anonymously challenge others using their question collections until their opponent provides an incorrect answer. With this revised structure, points serve as the initial incentive to engage with the Quiz feature. However, the interaction with others serves as an intrinsic motivator, encouraging users to strive for higher-difficulty questions to challenge others. Additionally, users may be motivated to continue learning or recalling information to avoid losing in these interactions.

Applying a similar rework infused with deeper game design insights could greatly enhance the experience for prospective users. However, before this step, gathering additional data on the project's reception and soliciting input from more citizens would be invaluable. This information will help define the direction for implementing gamification strategies to encourage blood donation participation.

Prototype Implementation Usability and Usefulness

- RQ2. Will our initial prototype implementation of the design be considered usable and useful in its current iteration?

Based on the initial average SUS score of 70.91, it seems that our proposed implementation is progressing in the right direction in terms of usability. Additionally, participants did not report significant issues regarding how to use the main options of the app, as they rated the difficulty level closer to "Somewhat Easy." However, similar to the reception, we cannot draw definitive conclusions due to the small sample size and the potential influence of the Hawthorne effect. Moreover, the scores may have been positively biased due to the presence of an instruction manual and the support provided. Taking these factors into account, we directed our attention to the individual responses for more in-depth discussion.

Regarding difficulty, the activities with lower scores were those related to managing the characters (*evolving and selecting*), as well as the added survey functionality. From the comments, it appears that the functionality for upgrading the characters to their additional forms is not intuitive. The issue may stem from the fact that the options are spread across different screens, making it challenging to locate and connect them. Consolidating all the actions related to character management onto a single screen, separate from the character acquisition process, could potentially make the interface easier to use. In the case of the survey, the only complaint received was regarding its length, with participants finding it too long to complete. However, it received the lowest rating among the activities, suggesting that other participants may have also encountered issues with it. We can hypothesize that, aside from the length of the activity itself, participants may have been dissatisfied with its mandatory status

XSL•FO
RenderX

rather than being optional. We could enhance the data collection process within the app by integrating it with gamification concepts, offering initial extrinsic rewards to users interested in participating. Ideally, we should also establish a loop that fosters participation through intrinsic motivation. This could involve designing activities or incentives that align with users' intrinsic interests, values, or desires for personal growth or contribution. Furthermore, from specific results of the SUS score, the participant who gave the lowest score (30) cited issues with the user interface. Taking this into consideration, future implementations of the proposed design should allocate adequate time for interface functionality and compatibility tests.

Another point for analysis from the SUS results is the average score assigned to item 1 ("I think that I would like to use this application frequently") of the survey. Although participants expressed positive sentiments regarding downloading the app and recommending it to others, the responses to item 1 indicated a nearly neutral position regarding the desire to use the app frequently. Indeed, the variation in results could stem from differing perspectives among user types. Nondonors might not use the app as frequently, even if they appreciate its concept. Similarly, donors with no deferral experience might use it for reference purposes, but perhaps not as frequently as deferred donors. Another possible and simple reason could be that participants might have had different interpretations of the term "frequently." Besides the inclusion of the "user category" variable, future survey evaluations could use a support question to help identify the regularity (if either daily, weekly, or monthly, as examples) of usage of our proposed implementation.

Further data collection is still required to obtain more detailed feedback about the current implementation, as there might be additional issues or shortcomings from the usability or the usefulness that were not captured because of the small number of participants.

Limitations

The study has multiple limitations that affect the reliability and generalization of the results. The small sample size of only 11 participants from Japan limits our ability to capture the true opinions of the various groups within the target population (prospective blood donors). Nondonors, donors, and deferred donors could have different perspectives and specific improvements regarding the design. Besides, although the recruitment was performed with Japanese material, we cannot confirm that only Japanese citizens participated in the evaluation. We have to consider that, although blood donation is seen as an altruistic and social activity in general, there can be differences in how individual values contribute to society according to one's cultural background.

Recruiting participants through social networks and including ownership of an iPhone as part of the criteria may have biased the sample toward individuals with higher levels of technological literacy, potentially influencing the results of the SUS score. However, to mitigate this bias, consistent guidance materials and tasks were provided to ensure a similar starting point for all participants.

The final version of the prototype for evaluation was created within a limited timeframe and programmed solely by 1 (REC) researcher. This constraint impacted the resources available for implementing content and graphical user interface options. The workforce constraints also impacted the choice of the target system for development, leading to the selection of iOS for release due to the developer's familiarity with it. Additionally, while the introduction of elements and activities involving user donation was considered, the acquisition and integration of these data into the current iteration of the project proved infeasible due to limitations related to permissions, partnerships, and time constraints.

Not collecting quantitative results regarding the value of the gamification aspect of the proposed design represents a significant weakness in the evaluation. Furthermore, the anonymous nature of the responses prevented the possibility of soliciting more detailed explanations regarding certain qualitative answers or comments from participants about the gamified components of the app. Indeed, it is crucial to address these shortcomings in future evaluations. Planning for a recruitment process that ensures a sufficient number of participants and obtaining ethical approval are essential steps for conducting a more comprehensive evaluation.

Conclusions

ICT systems have gained significant recognition and reliability across various fields, including within the realm of blood donation. We sought to explore previous work related to deferred donors and identify areas for further improvement. In addition to providing automated services, certain ICT projects have prioritized enhancing user motivation by incorporating gamification into their design. However, upon reviewing the current literature, it became apparent that only a few, if any, of the existing systems have specifically addressed the experience of deferral or its implications. In this research, we introduced an innovative ICT gamified design and implementation aimed at addressing this overlooked issue. Additionally, we offered an initial assessment of the project's potential reception, usability, and usefulness. Further enhancements can be made to the design of activities, which currently rely primarily on extrinsic motivation elements, to incorporate more social interaction. This would create an enriched activity loop that fosters intrinsic motivation. Further research could involve a more specialized and longitudinal design evaluation with a larger sample size. Understanding which specific features or gamification elements influence citizens' intentions or behaviors regarding their role in blood donation could be crucial for future design endeavors. Moreover, it could serve as a reference point for official ICT implementations in blood donation services.

XSL•FO
RenderX

Acknowledgments

The authors thank the citizens of Kyoto for their participation and feedback regarding the prototype of the study. REC holds a Japanese Government (Monbukagakusho: MEXT) Scholarship for the Doctoral Program. This work was supported by JST (CREST grant JPMJCR21M1), Japan.

Authors' Contributions

REC performed the literature research, design of the study, development of the app, and data analysis. LHOS and YM reviewed and validated the steps of the study design, especially the preliminary evaluation of the concept. All authors contributed points for discussion and implications based on the results. Additionally, all authors participated in drafting and revising the manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Study recruitment flyer—Japanese version.
[PNG File , 767 KB - humanfactors_v11i1e50086_app1.png]

Multimedia Appendix 2
Informed consent—Google Form sections for recruitment.
[DOCX File , 45 KB - humanfactors_v11i1e50086_app2.docx]

Multimedia Appendix 3
App user manual—English version (from right to left).
[PDF File (Adobe PDF File), 2982 KB - humanfactors_v11i1e50086_app3.pdf]

Multimedia Appendix 4
List of tasks for the app evaluation.
[DOCX File , 20 KB - humanfactors_v11i1e50086_app4.docx]

Multimedia Appendix 5
Study main survey—acceptance, System Usability Scale (SUS), and feedback.
[DOCX File , 22 KB - humanfactors_v11i1e50086_app5.docx]

Multimedia Appendix 6
Study survey data—formatted, anonymized, and translated to English.
[XLSX File (Microsoft Excel File), 17 KB - humanfactors_v11i1e50086_app6.xlsx]

Multimedia Appendix 7
System Usability Scale (SUS) results from each user with average and SD.
[PNG File , 150 KB - humanfactors_v11i1e50086_app7.png]

References

1. Gemelli CN, Thijsen A, Van Dyke N, Masser BM, Davison TE. Emotions experienced when receiving a temporary deferral: perspectives from staff and donors. ISBT Science Series 2018 Oct 14;13(4):394-404. [doi: 10.1111/voxs.12463]
2. van Dongen A. Easy come, easy go. Retention of blood donors. Transfus Med 2015 Aug 07;25(4):227-233. [doi: 10.1111/tme.12249] [Medline: 26399971]
3. Spekman ML, van Tilburg TG, Merz E. Do deferred donors continue their donations? A large-scale register study on whole blood donor return in the Netherlands. Transfusion 2019 Dec 17;59(12):3657-3665 [FREE Full text] [doi: 10.1111/trf.15551] [Medline: 31621923]
4. Davison TE, Masser BM, Gemelli CN. Deferred and deterred: a review of literature on the impact of deferrals on blood donors. ISBT Science Series 2019 Aug 28;15(1):3-10. [doi: 10.1111/voxs.12520]
5. Tynell E, Norda R, Ekermo B, Sanner M, Andersson S, Björkman A. False-reactive microbiologic screening test results in Swedish blood donors-how big is the problem? A survey among blood centers and deferred donors. Transfusion 2007 Jan;47(1):80-89. [doi: 10.1111/j.1537-2995.2007.01067.x] [Medline: 17207234]

XSL•FO
RenderX

6. Mathew S, King M, Glynn S, Dietz S, Caswell S, Schreiber G. Opinions about donating blood among those who never gave and those who stopped: a focus group assessment. Transfusion 2007 Apr;47(4):729-735. [doi: 10.1111/j.1537-2995.2007.01177.x] [Medline: 17381633]
7. Eder A, Goldman M, Rossmann S, Waxman D, Bianco C. Selection criteria to protect the blood donor in North America and Europe: past (dogma), present (evidence), and future (hemovigilance). Transfus Med Rev 2009 Jul;23(3):205-220. [doi: 10.1016/j.tmrv.2009.03.003] [Medline: 19539875]
8. Clement M, Shehu E, Chandler T. The impact of temporary deferrals on future blood donation behavior across the donor life cycle. Transfusion 2021 Jun 29;61(6):1799-1808. [doi: 10.1111/trf.16387] [Medline: 33780022]
9. Rahimi B, Nadri H, Lotfnezhad Afshar H, Timpka T. A Systematic Review of the Technology Acceptance Model in Health Informatics. Appl Clin Inform 2018 Jul 15;9(3):604-634 [FREE Full text] [doi: 10.1055/s-0038-1668091] [Medline: 30112741]
10. Morales J, Inupakutika D, Kaghyan S, Akopian D, Yin Z, Evans M, et al. Technology-based health promotion: Current state and perspectives in emerging gig economy. Biocybern Biomed Eng 2019 Jul;39(3):825-842 [FREE Full text] [doi: 10.1016/j.bbe.2019.07.006] [Medline: 32313347]
11. Sundarde C, Jain S, Shaikh E. Advancement of Blood Donation Application. Academia. 2015 Nov. URL: https://www.academia.edu/19419947/Advancement_of_Blood_Donation_Application [accessed 2022-01-21]
12. Ouhbi S, Fernández-Alemán JL, Toval A, Idri A, Pozo JR. Free blood donation mobile applications. J Med Syst 2015 May 3;39(5):52. [doi: 10.1007/s10916-015-0228-0] [Medline: 25732077]
13. Hillgrove TL, Doherty KV, Moore VM. Understanding non-return after a temporary deferral from giving blood: a qualitative study. BMC Public Health 2012 Dec 10;12(1):1063 [FREE Full text] [doi: 10.1186/1471-2458-12-1063] [Medline: 23227817]
14. Deterding S, Sicart M, Nacke L, O'Hara K, Dixon D. Gamification. using game-design elements in non-gaming contexts. In: CHI EA '11: CHI '11 Extended Abstracts on Human Factors in Computing Systems. New York, United States: Association for Computing Machinery; 2011 May 07 Presented at: CHI '11: CHI Conference on Human Factors in Computing Systems; May 7 - 12, 2011; Vancouver, Canada p. 2425-2428 URL: https://dl.acm.org/doi/10.1145/1979742.1979575 [doi: 10.1145/1979742.1979575]
15. Huotari K, Hamari J. Defining gamification: a service marketing perspective. In: MindTrek '12: Proceeding of the 16th International Academic MindTrek Conference. New York, United States: Association for Computing Machinery; 2012 Oct 03 Presented at: AcademicMindTrek '12: International Conference on Media of the Future; 2012; Tampere Finland p. 17-22 URL: https://dl.acm.org/doi/10.1145/2393132.2393137 [doi: 10.1145/2393132.2393137]
16. Domingos DCL, Lima LFSG, Messias TF, Feijo JVL, Diniz AAR, Soares HB. Blood hero: An application for encouraging the blood donation by applying gamification. : IEEE; 2016 Aug Presented at: 38th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC'16); August 16 - 20, 2016; Orlando, FL, USA p. 5624-5627. [doi: 10.1109/EMBC.2016.7592002]
17. Sabani AC, Manuaba IBK, Adi E. Gamification: Blood Donor Apps for IOS Devices. Journal of Games, Game Art, and Gamification 2021 Oct 19;1(1):14-26. [doi: 10.21512/jggag.v1i1.7247]
18. Xu L, Shi H, Shen M, Ni Y, Zhang X, Pang Y, et al. The Effects of mHealth-Based Gamification Interventions on Participation in Physical Activity: Systematic Review. JMIR Mhealth Uhealth 2022 Feb 03;10(2):e27794 [FREE Full text] [doi: 10.2196/27794] [Medline: 35113034]
19. Ossenbrink L, Haase T, Timpel P, Schoffer O, Scheibe M, Schmitt J, et al. Effectiveness of Digital Health Interventions Containing Game Components for the Self-management of Type 2 Diabetes: Systematic Review. JMIR Serious Games 2023 Jun 01;11:e44132 [FREE Full text] [doi: 10.2196/44132] [Medline: 37261900]
20. Ferguson E, France CR, Abraham C, Ditto B, Sheeran P. Improving blood donor recruitment and retention: integrating theoretical advances from social and behavioral science research agendas. Transfusion 2007 Nov;47(11):1999-2010. [doi: 10.1111/j.1537-2995.2007.01423.x] [Medline: 17958528]
21. Giles M, Cairns E. Blood donation and Ajzen's theory of planned behaviour: an examination of perceived behavioural control. Br J Soc Psychol 1995 Jun 06;34 (Pt 2)(2):173-188. [doi: 10.1111/j.2044-8309.1995.tb01056.x] [Medline: 7620844]
22. Khani Jeihooni A, Jormand H, Saadat N, Hatami M, Abdul Manaf R, Afzali Harsini P. The application of the theory of planned behavior to nutritional behaviors related to cardiovascular disease among the women. BMC Cardiovasc Disord 2021 Dec 07;21(1):589 [FREE Full text] [doi: 10.1186/s12872-021-02399-3] [Medline: 34876014]
23. Conner M, Armitage CJ. Extending the Theory of Planned Behavior: A Review and Avenues for Further Research. J Applied Social Pyschol 2006 Jul 31;28(15):1429-1464. [doi: 10.1111/j.1559-1816.1998.tb01685.x]
24. Robinson NG, Masser BM, White KM, Hyde MK, Terry DJ. Predicting intentions to donate blood among nondonors in Australia: an extended theory of planned behavior. Transfusion 2008 Dec 05;48(12):2559-2567. [doi: 10.1111/j.1537-2995.2008.01904.x] [Medline: 18717776]
25. Masser BM, Bednall TC, White KM, Terry D. Predicting the retention of first-time donors using an extended Theory of Planned Behavior. Transfusion 2012 Jun 18;52(6):1303-1310 [FREE Full text] [doi: 10.1111/j.1537-2995.2011.03479.x] [Medline: 22257135]
26. Masser BM, White KM, Hyde MK, Terry DJ. The psychology of blood donation: current research and future directions. Transfus Med Rev 2008 Jul;22(3):215-233. [doi: 10.1016/j.tmrv.2008.02.005] [Medline: 18572097]

XSL•FO
RenderX

27. Bagot KL, Murray AL, Masser BM. How Can We Improve Retention of the First-Time Donor? A Systematic Review of the Current Evidence. Transfus Med Rev 2016 Apr;30(2):81-91. [doi: 10.1016/j.tmrv.2016.02.002] [Medline: 26971186]
28. Bednall TC, Bove LL, Cheetham A, Murray AL. A systematic review and meta-analysis of antecedents of blood donation behavior and intentions. Soc Sci Med 2013 Nov;96:86-94. [doi: 10.1016/j.socscimed.2013.07.022] [Medline: 24034955]
29. Aarts H, Verplanken B, van Knippenberg A. Predicting Behavior From Actions in the Past: Repeated Decision Making or a Matter of Habit? J Applied Social Pyschol 2006 Jul 31;28(15):1355-1374. [doi: 10.1111/j.1559-1816.1998.tb01681.x]
30. Callero PL. Role-Identity Salience. Social Psychology Quarterly 1985 Sep;48(3):203. [doi: 10.2307/3033681]
31. Charng H, Piliavin JA, Callero PL. Role Identity and Reasoned Action in the Prediction of Repeated Behavior. Social Psychology Quarterly 1988 Dec;51(4):303. [doi: 10.2307/2786758]
32. Hagger MS, Chatzisarantis NLD. Integrating the theory of planned behaviour and self‐determination theory in health behaviour: A meta‐analysis. British J Health Psychol 2010 Dec 24;14(2):275-302. [doi: 10.1348/135910708x373959]
33. Williams LA, Sun J, Masser B. Integrating self‐determination theory and the theory of planned behaviour to predict intention to donate blood. Transfusion Med 2018 Nov 11;29(S1):59-64. [doi: 10.1111/tme.12566]
34. Pentecost R, Arli D, Thiele S. It's my choice! Investigating barriers to pro-social blood donating behaviour. MIP 2017 Apr 03;35(2):243-258. [doi: 10.1108/mip-03-2016-0055]
35. Deci EL. Self-determination theory. In: Ryan RM, editor. Handbook of Theories of Social Psychology. London: Sage Publications Ltd; 2011:416-436.
36. Ryan RM, Deci EL. Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. American Psychologist 2000;55(1):68-78. [doi: 10.1037/0003-066X.55.1.68]
37. France CR, France JL, Carlson BW, Frye V, Duffy L, Kessler DA, et al. Applying self-determination theory to the blood donation context: The blood donor competence, autonomy, and relatedness enhancement (Blood Donor CARE) trial. Contemp Clin Trials 2017 Feb;53:44-51 [FREE Full text] [doi: 10.1016/j.cct.2016.12.010] [Medline: 27979752]
38. M'Sallem W. Role of motivation in the return of blood donors: mediating roles of the socio-cognitive variables of the theory of planned behavior. Int Rev Public Nonprofit Mark 2021 Jun 23;19(1):153-166. [doi: 10.1007/s12208-021-00295-2]
39. DiTommaso D. Beyond gamification: Architecting engagement through game design thinking. Slideshare, a Scribd company. Nov 11. URL: https://www.slideshare.net/slideshow/beyond-gamification-architecting-engagement-through-game-design-thinking/10125282 [accessed 2024-02-13]
40. DiTommaso D, Taylor C. Beyond gamification: designing behavior change games. In: CHI PLAY '14: Proceedings of the first ACM SIGCHI annual symposium on Computer-human interaction in play. New York, United States: Association for Computing Machinery; 2014 Oct 19 Presented at: CHI PLAY '14: The annual symposium on Computer-Human Interaction in Play; October 19 - 21, 2014; Toronto Ontario Canada p. 475 URL: https://dl.acm.org/doi/10.1145/2658537.2662410 [doi: 10.1145/2658537.2662410]
41. Hunicke R, LeBlanc M, Zubek R. MDA: A Formal Approach to Game Design and Game Research. In: AAAI Workshop Papers 2004. 2004 Presented at: Nineteenth National Conference on Artificial Intelligence (AAAI-04); July 25 - 29, 2004; San Jose, California p. A.
42. Ryan, Deci. Intrinsic and Extrinsic Motivations: Classic Definitions and New Directions. Contemp Educ Psychol 2000 Jan;25(1):54-67. [doi: 10.1006/ceps.1999.1020] [Medline: 10620381]
43. Hassan L. Governments Should Play Games. Simulation & Gaming 2016 Dec 19;48(2):249-267. [doi: 10.1177/1046878116683581]
44. Koivisto J, Hamari J. The rise of motivational information systems: A review of gamification research. International Journal of Information Management 2019 Apr;45:191-210. [doi: 10.1016/j.ijinfomgt.2018.10.013]
45. Vallerand RJ, Blssonnette R. Intrinsic, Extrinsic, and Amotivational Styles as Predictors of Behavior: A Prospective Study. J Personality 1992 Sep 28;60(3):599-620. [doi: 10.1111/j.1467-6494.1992.tb00922.x]
46. Ryan RM, Rigby CS, Przybylski A. The Motivational Pull of Video Games: A Self-Determination Theory Approach. Motiv Emot 2006 Nov 29;30(4):344-360. [doi: 10.1007/s11031-006-9051-8]
47. Marczewski A. Introduction to gamification part 4: motivation (R.A.M.P, Maslow, SDT and more). Gamified UK. 2019. URL: https://www.gamified.uk/2019/01/30/introduction-to-gamification-part-4-motivation-r-a-m-p-maslow-sdt-and-more/ [accessed 2022-03-15]
48. Werbach K, Hunter D. For the Win: How Game Thinking Can Revolutionize Your Business. Philadelphia, PA: Wharton School Press; Oct 30, 2012.
49. Espinoza R, Hiragi S, Santos L, Morris K, Ueshima H, Yamamoto G, et al. Blood donation and deferral awareness in Japan: issues and opportunities for information and communications technology interventions. In: Proceedings of the Annual Conference of the Institute of Systems, Control and Information Engineers. 2021 May Presented at: The 65th Annual Conference of the Institute of Systems, Control and Information Engineers (SCI'21); May 26-28, 2021; Okayama University (Online) p. 1112-1116 URL: https://cir.nii.ac.jp/crid/1520853834864279168
50. Gemelli CN, Kruse SP, Thijsen A, Van Dyke N, Davison TE. Educating donors about deferrals through the provision of tailored materials: Findings from a survey of deferred donors. Transfusion 2020 Dec 07;61(3):822-829. [doi: 10.1111/trf.16212]
51. Takashi M. Irasutoya. URL: https://www.irasutoya.com [accessed 2021-03-13]

XSL•FO
RenderX

52. Sauro J, Lewis JR. Chapter 8: Standardized usability questionnaires. In: Quantifying the User experience: Practical Statistics for User Research. Burlington, MA: Morgan Kaufmann; 2012:185-240.
53. Brooke J. SUS: A 'Quick and Dirty' Usability Scale. In: Usability Evaluation In Industry. London, England: CRC Press; Nov 1996.
54. Teske K. Duolingo. CALICO 2017 Aug 24;34(3):393-401 [FREE Full text] [doi: 10.1558/cj.32509]
55. Atherley AE, Taylor CG, Whittington A, Jonker C. Knowledge, attitudes and practices towards blood donation in Barbados. Transfus Med 2016 Dec 16;26(6):415-421. [doi: 10.1111/tme.12359] [Medline: 27634655]
56. Adair JG. The Hawthorne effect: A reconsideration of the methodological artifact. Journal of Applied Psychology 1984;69(2):334-345. [doi: 10.1037//0021-9010.69.2.334]
57. Kapp KM. The Gamification of Learning and Instruction Fieldbook: Ideas into Practice. Hoboken, NJ: Wiley; 2013.

Abbreviations

ICT: information and communications technology
RQ: research question
SDT: Self-Determination Theory
SUS: System Usability Scale
TPB: Theory of Planned Behavior

Edited by A Kushniruk; submitted 01.07.23; peer-reviewed by F Gallego Duran, T Barros Pontes e Silva; comments to author 22.08.23; revised version received 31.10.23; accepted 07.04.24; published 14.06.24.

Please cite as:
Espinoza Chamorro R, Santos LHO, Mori Y, Liu C, Yamamoto G, Kuroda T
Gamification Approach to Provide Support About the Deferral Experience in Blood Donation: Design and Feasibility Study
JMIR Hum Factors 2024;11:e50086
URL: https://humanfactors.jmir.org/2024/1/e50086
doi:10.2196/50086
PMID:38875005

©Roberto Espinoza Chamorro, Luciano H O Santos, Yukiko Mori, Chang Liu, Goshiro Yamamoto, Tomohiro Kuroda. Originally published in JMIR Human Factors (https://humanfactors.jmir.org), 14.06.2024. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on https://humanfactors.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Interest in mHealth Among Patients With Low Back Pain: Cross-Sectional Study

Jonas Ammundsen Ipsen[1,2], MSc; Louise Fleng Sandal[3], PhD; Natalie Hong Siu Chang[1,4], MSc; Berit Schiøttz-Christensen[5], MD, PhD; Karen Søgaard[3,6], PhD; Anders Hansen[1,4], PhD

[1]Department of Regional Health Research, University of Southern Denmark, Odense, Denmark
[2]Department of Physical Therapy and Occupational Therapy, Lillebaelt Hospital, University Hospital of Southern Denmark, Kolding, Denmark
[3]Department of Sports Science and Clinical Biomechanics, University of Southern Denmark, Odense, Denmark
[4]Medical Research, Spine Centre of Southern Denmark, University Hospital of Southern Denmark, Middelfart, Denmark
[5]Research Unit of General Practice, University of Southern Denmark, Odense, Denmark
[6]Department of Clinical Research, University of Southern Denmark, Odense, Denmark

Corresponding Author:
Jonas Ammundsen Ipsen, MSc
Department of Physical Therapy and Occupational Therapy
Lillebaelt Hospital
University Hospital of Southern Denmark
Sygehusvej 24
Kolding, 6000
Denmark
Phone: 45 76362832
Email: jonas.ammundsen.ipsen@rsyd.dk

Abstract

Background: Digitally supported self-management tailored to an individual's need, in addition to usual care, may reduce pain-related disability compared to usual care alone, and patients with low back pain (LBP) using mobile health (mHealth) solutions express positive experiences. Hence, implementing mHealth solutions designed to support self-management is desirable from a clinical and patient perspective. Easily accessible mHealth solutions that can support the self-management of patients with LBP are available, but interest may be subgroup specific. Understanding the characteristics and preferences of patients with LBP labeled as interested may help to reach relevant LBP patient groups and inform the development and implementation of effective interventions with mHealth for patients with LBP.

Objective: This study aims to explore the proportion of patients with LBP labeled as interested in testing an mHealth solution designed to support self-management in addition to usual care and to assess how these patients differ from those who were labeled as not interested.

Methods: This exploratory cross-sectional study analyzed demographic and patient-reported outcomes from the SpineData registry, a Danish registry of patients with LBP in an outpatient setting. Between February and December 2019, the SpineData registry was used to assess the preliminary eligibility of patients for a clinical trial (selfBACK). Patients were labeled as interested or uninterested depending on if they responded to an invitation to be tested for eligibility for the trial Outcomes were selected from the International Classification of Functioning core set of LBP using a clinical approach. Associations were assessed in a backward selection process, and the proportion of variance explained was assessed with pseudo-R^2 statistic.

Results: This study included 843 patients, with 181 (21%) individuals labeled as interested in participating in the selfBACK trial. Notably, the cohort labeled as interested differed from their uninterested counterparts in two key aspects: age (36-65 years: 116/181, 64.1% vs 347/662, 52.4%; P=.003) and smoking status (smokers: 22/181, 12.5% vs 174/662, 26.6%; P<.001). Those aged 36-65 years had higher odds of being labeled as interested compared to individuals aged 18-35 years (odds ratio [OR] 0.43, 95% CI 0.26-0.71) and those 65 years or older (OR 0.77, 95% CI 0.53-1.15). Nevertheless, age accounted for only a modest proportion of variance (R^2=0.014). Smokers demonstrated lower odds of being labeled as interested (OR 0.39, 95% CI 0.24-0.64), with smoking status explaining a similarly small proportion of variance (R^2=0.019). Collectively, age and smoking status accounted for 3.3% of the variance.

XSL•FO
RenderX

Conclusions: Our investigation revealed that 181 (21%) individuals with LBP invited to participate in the mHealth solution trial for self-management expressed interest. Generally, the characteristics of those labeled as interested and uninterested were comparable. Of note, patients aged 36-65 years had a higher frequency of being labeled as interested compared to their younger and older counterparts.

(JMIR Hum Factors 2024;11:e48729) doi:10.2196/48729

KEYWORDS

low back pain; mHealth solutions; mobile health; characteristics; patient interest; transferability; representativeness

Introduction

Digital health interventions delivered with smartphones (mobile health [mHealth] solutions) are accessible to most patients across age, geography, and socioeconomic status. Thus, clinicians' expectations of mHealth solutions are significant, and the availability of new solutions on the commercial market every day also indicates a strong general interest in using mHealth solutions [1,2]. Nevertheless, many mHealth solutions have limited download rates, and if downloaded, the use can be scarce [3,4]. This discrepancy may indicate a need for a better understanding of potential users and their characteristics.

For patients with low back pain (LBP; not attributable to a recognizable, known specific pathology such as infections, fractures, or structural deformity), self-management support is recommended as the first line of treatment [5-8]. This may involve empowering patients to know when to consult for diagnostic assessment, symptom relief, or advice [9]. Digitally supported self-management may be delivered through smartphone apps or digital platforms to facilitate and enhance such self-management practices. Research indicates that the integration of such digitally supported self-management strategies, when combined with standard care, can lead to a reduction in pain-related disability [10]. Further, evidence supports that mHealth solutions designed to support self-management are accepted by patients with chronic LBP [9]. Therefore, there is a growing interest in implementing mHealth solutions designed to support self-management from both clinical and patient perspectives. However, despite the potential benefits, the level of acceptance and use of these interventions remains an area that requires further investigation.

However, studies on other patient groups using mHealth solutions report that lower age, higher education, higher income, higher BMI, and higher self-perceived health are associated with increased use [4,11]. In contrast, the cost of using these apps is a significant barrier [11].

Individuals with LBP who use mHealth solutions to self-manage may thus represent a specific subset within the general population. Therefore, this study aimed to investigate the percentage who expressed interest in participating in a trial evaluating an mHealth solution designed for self-management alongside standard care, as well as to evaluate potential distinctions between those who were labeled as expressing interest and those who were not.

Methods

Study Design

This exploratory cross-sectional study used demographic and patient-reported outcomes (PROs) from an internet-based multiuser clinical registry (SpineData) [12]. Reporting follows the STROBE (Strengthening the Reporting of Observational Studies in Epidemiology) guidelines [13].

Setting

Data were collected at the Spine Centre of Southern Denmark, an outpatient hospital that performs clinical spine evaluations [12]. General practitioners or chiropractors typically refer patients to the Spine Centre, which performs a multidisciplinary assessment of its patients, with more than 10,000 new cases yearly.

Before patients are evaluated at the Spine Centre, they provide information in the local SpineData registry [14]. The registry is designed based on the biopsychosocial model of health, and information is collected across the health domains of pain, activity limitation, work participation, psychological factors, physical impairment, and contextual factors [12]. To mitigate nonresponse and missing information, SpineData uses a "waterfall" model (eg, patients in employment are not asked to respond to causes for unemployment). SpineData has an overall completion rate of 80% and approximately 60% of patients agreed to their responses being used for research [14]. The use of this registry allows for the comprehensive assessment of patients consulted at the Spine Centre and provides a rich source of data for research studies, such as the one presented in this paper.

Participants

Between February and December 2019, SpineData was used to identify eligible patients based on the following criteria: consenting to be contacted for research projects, proficiency in Danish, and experiencing LBP in the past 14 days that exceeded their leg pain in severity. Patients with previous back surgery, who were actively filing for a pension, or who were younger than 18 years were not invited. All patients matching the eligibility criteria were sent a letter of invitation to hear more about the selfBACK trial. One reminder was sent. The patients who did not respond to either invitation or reminder were labeled uninterested. The selfBACK trial investigated the effectiveness of the selfBACK digital decision support system that provided patients with LBP individually tailored digital support in an app format using three content domains: (1) physical activity, (2) education, and (3) exercise programs. The trial investigated the

XSL•FO
RenderX

additive effect of the selfBACK system in addition to usual care. Participants in this trial were recruited from primary health care such as chiropractors, physiotherapists, and general practitioners in addition to the Spine Center of Southern Denmark. Recruitment was performed in Denmark and Norway. The population within this study concerns the pool of patients seen at the Spine Center, who would have received an invitation to eligibility screening to the selfBACK trial based on their answers given in the SpineData clinical registry. In this study, all patients who matched the preliminary eligibility criteria for the selfBACK trial were included [15].

Outcomes

The variables of interest were selected from the SpineData registry, based on the International Classification of Function core set for LBP and clinical reasoning [16]. The demographics and clinical characteristics comprised the domains of pain, activity limitation, work participation, and psychological and contextual factors (Textbox 1).

XSL•FO
RenderX

Textbox 1. Detailed description of the content and handling of included outcomes.

Sex

- Male or female

Age

- Patients were categorized into age groups ≤35, 36-65, and >65 years.

BMI

- The anthropometric variables of height and weight were used to calculate BMI (kg/m^2).

Smoker status

- Categorical variable that was dichotomized to smoker and nonsmoker strata. If a patient indicated cigarette use of any kind, they were categorized as a smoker.

Alcohol consumption

- Categorical variable that was stratified into two groups based on the consumption of more than 14 alcoholic beverages a week. The threshold was based on the recommendation of the Danish Ministry of Health [17].

Comorbidities

- This variable was based on four dichotomous variables: allergies, including medication; cancers; heart disease; and lung disease. If a patient replied yes to one of these variables, they were categorized as having comorbidities.

Current work status

- The work status variables consisted of different ways of participating in the labor market: working full- or part-time, flex job, in education, job training due to inability to maintain habitual job function, unemployed, early retirement, pensions, stay at home, and other. The variable was dichotomized to working or not by grouping patients indicating working part- or full-time, flex, and students in one group and the remainder in another group.

Multiple pain sites

- SpineData contains a freehanded pain drawing. The pain drawing was post defined into 46 anatomical regions. In this study, the regions were grouped into 9 areas: neck, shoulders, upper back, elbows, lower back, wrists/hands, hips/thighs, knees, and ankles/feet, inspired by Øverås et al [18]. Patients with two or more pain sites were considered as having multiple pain sites.

Average back pain

- The average back pain in the last 14 days was measured on a 0-10 numeric rating scale, with 10 indicating the worst imaginable pain.

STarT BaCK screening tool [19]

- The STarT Back scores categorize patients into three strata based on their risk of developing chronicity: low risk, moderate risk, and high risk of chronicity:
  - Low risk: <3
  - Moderate risk: ≥4 and subscore ≤3
  - High risk: ≥4 and subscore ≥4

EQ-5D-5L-VAS [20]

- Numeric rating scale score spanning from 0 to 100, with 100 representing the best possible health state

Oswestry Disability Index (ODI) [21]

- The ODI is a questionnaire containing 10 items that are scored from 0 to 5. The maximum score is 50 points, which indicates that the patient is bedbound. The ODI has been found valid for patients with low back pain [22].
- To estimate the patients' functional level, the ODI Stata package was used. The ODI package allows for the imputation of data for one missing value. The missing values in one section were replaced with the average score for all sections.

Anxiety [23]

- Numeric score rating from 0 to 10, with 0 indicating no anxiety and 10 a high degree of anxiety

XSL•FO
RenderX

Social isolation [23]

- Numeric score rating from 0 to 10, with 0 indicating no loneliness and 10 a high degree of loneliness

Catastrophization (terrible pain that will never improve) [23]

- Numeric score rating from 0 to 10, with 0 indicating no catastrophization and 10 a high degree of catastrophization

Catastrophization (the pain is overwhelming) [23]

- Numeric score rating from 0 to 10, with 0 indicating no catastrophization and 10 a high degree of catastrophization

Risk of persisting pain [23]

- Numeric score rating from 0 to 10, with 0 indicating no risk of persisting pain and 10 a high risk of persisting pain

Feelings of sadness, depression, or hopelessness [23]

- Numeric score rating from 0 to 10, with 0 indicating no feelings of depression and 10 a constant presence of depression

Loss of interest or joy [23]

- Numeric score rating from 0 to 10, with 0 indicating no loss of interest or joy and 10 never feeling interest or joy

Fearing activity will damage the back [23]

- Numeric score rating from 0 to 10, with 0 indicating no fear that physical activity will damage the back and 10 completely agreeing that physical activity will damage the back

Fearing activity will increase the pain [23]

- Numeric score rating from 0 to 10, with 0 indicating completely disagreeing to avoid physical activity and 10 completely agreeing to avoid physical activity

Exposure

Patients were allocated into two groups based on their response to being invited for eligibility screening for the selfBACK trial. Those who responded positively to the invitation to be screened were labeled as interested in using the mHealth solution, whereas those who did not respond were labeled as uninterested.

Statistical Methods

The demographics and baseline characteristics of patients who were or were not labeled as interested in the digital mHealth intervention were assessed using the χ^2 test for categorical variables and 2-tailed Student *t* test for continuous variables. Baseline characteristics are reported as the proportion and percentage or mean and SD.

To assess the strength of associations between PROs and patients labeled as interested in mHealth or not, we used univariate and multivariate logistic regression analysis with an odds ratio (OR) and 95% CI. The associations were assessed in a backward selection process, and the proportion of variance explained was assessed with McFadden pseudo-R^2 statistic. Statistical analyses were performed with Stata statistical software (Release 17; StataCorp LLC). Missing information was handled using pairwise deletion. The ODI Stata package allows for data imputation for one missing value. The missing values in one section were replaced with the average score for all sections. To avoid overparameterizing the model, we aimed for a 1:10 patient-to-variable ratio.

Ethical Considerations

The Region of Southern Denmark was the data controller for this project, which is included in its records on personal data processing activities (file 21/13433). Data processing in the project was regulated by the Danish Act on Research Ethics Review of Health Research Projects section 14, subsection 2, which states that health research based solely on questionnaire surveys and registry data is exempt from the obligation to notify the committees. Following the Danish Health Care Act, we obtained approval for using hospital record data for scientific purposes from the council of the Region of Southern Denmark (file 21/25588). After merging, analyses were run on pseudonymized data, and the results presented in this manuscript do not enable the identification of single data participants. Hence, following national laws, no additional informed consent was collected and no remuneration was offered to patients.

Results

Overview

From February to the end of December 2019, 5796 patients (~80% of those invited) completed the SpineData registry before their diagnostic assessment at the Spine Centre. Of the total sample, 843 (15%) were invited to the selfBACK trial. The mean age of the cohort was 52 (SD 16.2) years, with an even distribution of sexes (male: n=429, 50.1%), and a mean BMI of 27.5 kg/m^2.

XSL•FO
RenderX

Of the 843 patients invited to the eligibility screen for the trial, 181 (21%) accepted the invitation and were stratified into the group who were labeled as interested in the mHealth solution.

Of the 21 included variables, 8 had complete responses, and none of the remaining 13 variables had more than 2.5% missing responses.

Comparison of Patients Who Were Labeled as Interested and Uninterested in an mHealth Solution

Patients labeled as interested in using the mHealth solution were aged 36-65 years (*P*=.003) and had a lower proportion of smokers (*P*<.001) compared to the patients labeled as uninterested. The remaining variables were not different between the patients labeled as interested and uninterested (Table 1).

Table 1. Baseline characteristics of patients labeled as interested in the mobile health solution compared to the uninterested patients.

Baseline characteristic[a]	Interested (n=181)	Uninterested (n=662)	*P* value
Female, n (%)	93 (51.3)	321 (48.5)	.49
Age (years), n (%)			.003
18-35	21 (11.6)	146 (22.1)	
36-65	116 (64.1)	347 (52.4)	
>65	44 (24.3)	169 (25.5)	
BMI (kg/m^2), mean (SD)	28.1 (5.8)	27.2 (5.0)	.05
Smokers, n (%)	22 (12.5)	174 (26.6)	<.001
<14 alcohol consumption per week, n (%)	173 (95.5)	636 (96.0)	.76
Has comorbidities, n (%)	84 (46.7)	208 (42.3)	.29
Working, n (%)	95 (52.4)	367 (55.4)	.49
Has multiple pain sites, n (%)	136 (75.1)	462 (69.7)	.24
Average back pain (score range: 0-10), mean (SD)	6.3 (2.0)	6.3 (1.9)	.91
STarT Back tool, n (%)			.34
Low risk	54 (29.8)	164 (24.7)	
Moderate risk	46 (25.5)	169 (25.5)	
High risk	81 (44.7)	329 (49.7)	
EQ-5D-5L-VAS (score range: 0-100), mean (SD)	59.0 (22.1)	55.2 (23.0)	.05
Oswestry Disability Index (score range: 0-50), mean (SD)	30.3 (15.6)	31.1 (14.9)	.50
Anxiety (score range: 0-10), mean (SD)	3.8 (3.0)	3.8 (3.1)	.99
Loneliness (score range: 0-10), mean (SD)	1.4 (2.4)	1.3 (2.2)	.67
Catastrophization (terrible pain that will never improve; score range: 0-10), mean (SD)	4.8 (2.9)	5.0 (3.0)	.48
Catastrophization (the pain is overwhelming; score range: 0-10), mean (SD)	3.7 (3.1)	4.1 (3.1)	.24
Risk of persisting pain (score range: 0-10), mean (SD)	6.8 (2.6)	6.8 (2.6)	.91
Sadness (score range: 0-10), mean (SD)	3.5 (3.1)	3.6 (3.1)	.71
Loss of interest or joy (score range: 0-10), mean (SD)	4.3 (3.3)	4.3 (3.2)	.88
Fearing activity will damage the back (score range: 0-10), mean (SD)	3.4 (2.9)	3.8 (3.2)	.08
Fearing activity will increase the pain (score range: 0-10), mean (SD)	4.8 (3.2)	4.4 (3.3)	.16

[a]Missing: 8 of the 21 variables had complete responses, and none of the remaining 13 variables had more than 2.5% missing responses.

Our results suggest that patients aged 36-65 years were more likely to be labeled as interested in mHealth solutions compared to patients between 18-35 years (OR 0.43, 95% CI 0.026-0.711) and 65 years or older (OR 0.77, 95% CI 0.525-1.153) and explained a limited proportion of variance (R^2=0.014). Smoker (OR 0.39, 95% CI 0.244-0.636) and the association explained a limited proportion of variance (R^2=0.019). Combined, the associations of age and smoking explained 3.3% of the proportion of variance.

These findings were supported by univariate regression analysis and a comparison of patients who were labeled as expressing interest in the mHealth solution to those who did not. The proportion of variance explained in the group of patients labeled as interested in mHealth solutions across the 21 selected

XSL•FO
RenderX

variables was 0.059, with age and smoking status accounting for 0.033 of the variance (Table 2).

BMI (*P*=.05), overall perception of health measured using the EQ-5D-5L-VAS score (*P*=.05), and fear that activity will damage the back (*P*=.08) were borderline significant.

Table 2. Associations to being labeled as interested and proportion of variance explained.

	Odds ratio (95% CI)	SE	*Z*	People invited, N	*P* value (*P*>|z|)	R^2
Age (years)				830		0.014
36-65 (reference)	—	—	—		—	
18-35	0.43 (0.02-0.71)	0.114	–3.17		.002	
>65	0.77 (0.52-1.15)	0.151	–1.48		.14	
Smoking				830		0.019
Nonsmoker (reference)	—	—	—	—	—	—
Smoker	0.39 (0.24-0.64)	0.098	–3.72		<.001	

Discussion

Principal Results

This study aimed to explore the proportion of patients with LBP who were labeled as interested in using an mHealth solution designed to support self-management in addition to usual care and assess how these patients differed from those who were labeled as not interested. We found that 21% of the eligible patients were labeled as interested in using the mHealth solution. The groups had no statistically significant differences except that patients labeled as interested were more frequently within the 36-65 years age range and were nonsmokers.

Comparison With Prior Work

Previous evidence of the characteristics and associations of patients with LBP and their interest in mHealth solutions is limited. Contrary to Krebs and Duncan [4], we found a nonsignificant association between BMI and no association between being younger and labeled as interested in mHealth solutions. The key differences between Krebs and Duncan [4] and this study are the target populations (general population) and the type of mHealth solutions included (fitness apps or calorie trackers). Similar to our results, Philip et al [24] identified an association between higher age and increased use of mHealth solutions among patients with chronic pain. We suggest that the differences in results between Krebs and Duncan [4], Philip et al [24], and this study were due to differences between participants from the general population and patients with LBP or chronic pain. Three recent studies have assessed the characteristics and associations of users and nonusers of different mHealth apps, all using participants from the general population, but still lacking consensus. Walrave et al [25] identified no sociodemographic differences between users and nonusers of contact tracking alert apps, including the Belgian Corona alert app. A study of the general US population identified strong associations of age, gender, and education level with the use of fitness apps and calorie counters [26]. Lim et al [27] identified that female patients with higher education were more prevalent users of mHealth apps. Although this lack of consensus regarding patient interest could indicate a call for more research, it could also reflect that the interest in mHealth solutions may be characterized by patients' preferences and perspectives on the relevance of mHealth solutions.

Strengths and Limitations

This study benefitted from several strengths. First, we had access to comprehensive information on the patients participating through the SpineData registry. Further, we benefitted from the fact that SpineData has been in routine use for several years and is frequently updated per clinician and evidence demand [14]. Thus, the PROs were collected using validated questionnaires or questions designed for the LBP population and International Classification of Function core set [14,16,28,29]. The included patients were identified using a computer algorithm, and patients were sent one invitation and one reminder invitation to be screened for eligibility. Thus, the risk of unconscious bias in the recruitment was eliminated. However, using a single data source (SpineData) also limited the variables available to investigate in the univariate model. Low education and economic status have been associated with limited use and adoption of mHealth solutions [26,30], but this information was unavailable in SpineData. Smoking is reportedly more prevalent among patients with a lower socioeconomic or sociodemographic status [31,32]. Further, the use of one registry meant we only had access to PROs, which may be affected by recall bias. The statistically significant difference between being labeled as interested in mHealth solutions by smoking status could reflect a difference in education level. Thus, education level is a parameter that could differentiate the patients labeled as interested and those labeled as uninterested in the mHealth solution, although this hypothesis remains unanswered. Patients referred to the Spine Centre usually have pain for extended periods and at a higher intensity than patients in the primary sector [33]. Thus, these patients potentially have more complex LBP issues than those with LBP who were not referred, which means that our study population may be a subgroup of the general LBP population. The terms "interested" and "uninterested" pose a challenge due to their vague nature. We recognize the distinction between demonstrating a "cursory" interest and moving toward actual participation. After extensive discussions among authors, we chose the terms "interested" and "uninterested." Despite their less-than-optimal nature, we believe these terms best suit the context where we categorize patients based on their response to an invitation, progressing

XSL•FO
RenderX

from screening to eligibility for participation in a trial evaluating an mHealth solution supporting self-management in patients with LBP. Further, some patients might be interested in testing an mHealth solution but uninterested in participating in a trial or vice versa. Further, those labeled as uninterested in the mHealth solution in this study might see advantages in mHealth solutions that they found more relevant like how to stop smoking or lose weight [34]. This study only addresses patient characteristics; however, investigating clinicians' perspectives on the use and adoption of mHealth solutions in LBP self-management will similarly inform on barriers to and facilitators of increased mHealth adoption in clinical practice. However, as the SpineData clinical registry only entails patient data, this perspective was not possible in this study. Thus, the results of this study should be interpreted with caution regarding generalizability, and future qualitative or mixed methods studies could explore patients' preferences and perceptions of the relevance of mHealth solutions. Another important area of research can be clinicians' acceptability of mHealth solutions and the need for rigorous demonstrations of safety and efficacy to alleviate any reservations or hesitance among clinicians.

Conclusion

This study aimed to explore the characteristics of patients labeled as interested or uninterested in participating in a trial testing an mHealth solution designed to improve self-management. Our study identified that 21% (n=181) of eligible patients with LBP were labeled as interested in participating in the trial testing an mHealth solution to support self-management. Overall, the patients labeled as interested and uninterested, except for age and smoking status were similar. Interestingly, patients aged 36-65 years were more frequently labeled as interested in the mHealth solution. Thus, patients aged 36-65 years may be more interested in adopting mHealth solutions. How to increase interest in mHealth solutions among younger and older patients with LBP is an important consideration for future research and developers, especially as the findings of the selfBACK trial indicate an increased effect for older patients.

Authors' Contributions

AH, BSC, and KS conceptualized the study. AH, LFS, and NHSC acquired the data and data permissions. JAI performed the data analysis and drafted the manuscript, with support from NHSC, AH, and LFS. All authors helped draft and critically revised the manuscript for important intellectual content and approved the final version.

Conflicts of Interest

None declared.

References

1. Bates DW, Landman A, Levine DM. Health Apps and Health Policy: What Is Needed? JAMA 2018 Nov 20;320(19):1975-1976. [doi: 10.1001/jama.2018.14378] [Medline: 30326025]
2. Aitken M. Digital health trends 2021: innovation, evidence, regulation, and adoption. IQVIA. 2021 Jul 22. URL: https://www.iqvia.com/insights/the-iqvia-institute/reports-and-publications/reports/digital-health-trends-2021 [accessed 2023-11-15]
3. mHealth economics 2017/2018 – connectivity in digital health. Research2Guidance. 2017. URL: https://research2guidance.com/product/connectivity-in-digital-health/ [accessed 2022-12-25]
4. Krebs P, Duncan DT. Health app use among US mobile phone owners: a national survey. JMIR Mhealth Uhealth 2015 Nov 04;3(4):e101 [FREE Full text] [doi: 10.2196/mhealth.4924] [Medline: 26537656]
5. Hutting N, Johnston V, Staal JB, Heerkens YF. Promoting the use of self-management strategies for people with persistent musculoskeletal disorders: the role of physical therapists. J Orthop Sports Phys Ther 2019 Apr;49(4):212-215. [doi: 10.2519/jospt.2019.0605] [Medline: 30931733]
6. Caneiro JP, Roos EM, Barton CJ, O'Sullivan K, Kent P, Lin I, et al. It is time to move beyond 'body region silos' to manage musculoskeletal pain: five actions to change clinical practice. Br J Sports Med 2020 Apr;54(8):438-439. [doi: 10.1136/bjsports-2018-100488] [Medline: 31604698]
7. Corp N, Mansell G, Stynes S, Wynne-Jones G, Morsø L, Hill JC, et al. Evidence-based treatment recommendations for neck and low back pain across Europe: A systematic review of guidelines. Eur J Pain 2021 Feb;25(2):275-295 [FREE Full text] [doi: 10.1002/ejp.1679] [Medline: 33064878]
8. Qaseem A, Wilt TJ, McLean RM, Forciea MA, Denberg TD, Barry MJ, et al. Noninvasive treatments for acute, subacute, and chronic low back pain: a clinical practice guideline from the American College of Physicians. Ann Intern Med 2017 Apr 04;166(7):514-530 [FREE Full text] [doi: 10.7326/M16-2367] [Medline: 28192789]
9. Kongsted A, Ris I, Kjaer P, Hartvigsen J. Self-management at the core of back pain care: 10 key points for clinicians. Braz J Phys Ther 2021;25(4):396-406 [FREE Full text] [doi: 10.1016/j.bjpt.2021.05.002] [Medline: 34116904]
10. Sandal LF, Bach K, Øverås CK, Svendsen MJ, Dalager T, Stejnicher Drongstrup Jensen J, et al. Effectiveness of app-delivered, tailored self-management support for adults with lower back pain-related disability: a selfBACK randomized clinical trial. JAMA Intern Med 2021 Oct 01;181(10):1288-1296 [FREE Full text] [doi: 10.1001/jamainternmed.2021.4097] [Medline: 34338710]

XSL•FO
RenderX

11. Rai A, Chen L, Pye J, Baird A. Understanding determinants of consumer mobile health usage intentions, assimilation, and channel preferences. J Med Internet Res 2013 Aug 02;15(8):e149 [FREE Full text] [doi: 10.2196/jmir.2635] [Medline: 23912839]
12. Kent P, Kongsted A, Jensen TS, Albert HB, Schiøttz-Christensen B, Manniche C. SpineData - a Danish clinical registry of people with chronic back pain. Clin Epidemiol 2015;7:369-380 [FREE Full text] [doi: 10.2147/CLEP.S83830] [Medline: 26316820]
13. von Elm E, Altman DG, Egger M, Pocock SJ, Gøtzsche PC, Vandenbroucke JP, STROBE Initiative. The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies. Lancet 2007 Oct 20;370(9596):1453-1457 [FREE Full text] [doi: 10.1016/S0140-6736(07)61602-X] [Medline: 18064739]
14. Chang N. Rygmedicinsk Forskning - SpineData og MiRD databaserne. Sygehus Lillebælts. 2019. URL: http://tinyurl.com/48vu9h2m [accessed 2020-07-05]
15. Sandal LF, Stochkendahl MJ, Svendsen MJ, Wood K, Øverås CK, Nordstoga AL, et al. An App-Delivered Self-Management Program for People With Low Back Pain: Protocol for the selfBACK Randomized Controlled Trial. JMIR Res Protoc 2019 Dec 03;8(12):e14720 [FREE Full text] [doi: 10.2196/14720] [Medline: 31793897]
16. Kirschneck M, Kirchberger I, Amann E, Cieza A. Validation of the comprehensive ICF core set for low back pain: the perspective of physical therapists. Man Ther 2011 Aug;16(4):364-372. [doi: 10.1016/j.math.2010.12.011] [Medline: 21269870]
17. Notat: Sundhedsstyrelsens udmeldinger om indtag af alkohol. Sundhedsstyrelsen. 2023 Jul 19. URL: https://www.sst.dk/da/Viden/Alkohol/Alkoholforebyggelse/Sundhedsstyrelsens-udmeldinger-om-alkohol [accessed 2020-03-25]
18. Øverås CK, Nilsen TIL, Nicholl BI, Rughani G, Wood K, Søgaard K, et al. Multimorbidity and co-occurring musculoskeletal pain do not modify the effect of the SELFBACK app on low back pain-related disability. BMC Med 2022 Feb 08;20(1):53 [FREE Full text] [doi: 10.1186/s12916-022-02237-z] [Medline: 35130898]
19. STarTBack online. STarTBack. URL: https://startback.hfac.keele.ac.uk/training/resources/startback-online/ [accessed 2024-02-01]
20. Soer R, Reneman MF, Speijer BLGN, Coppes MH, Vroomen PCAJ. Clinimetric properties of the EuroQol-5D in patients with chronic low back pain. Spine J 2012 Nov;12(11):1035-1039. [doi: 10.1016/j.spinee.2012.10.030] [Medline: 23199409]
21. Fairbank JC, Pynsent PB. The Oswestry Disability Index. Spine (Phila Pa 1976) 2000 Nov 15;25(22):2940-52; discussion 2952. [doi: 10.1097/00007632-200011150-00017] [Medline: 11074683]
22. Schwind J, Learman K, O'Halloran B, Showalter C, Cook C. Different minimally important clinical difference (MCID) scores lead to different clinical prediction rules for the Oswestry disability index for the same sample of patients. J Man Manip Ther 2013 May;21(2):71-78 [FREE Full text] [doi: 10.1179/2042618613Y.0000000028] [Medline: 24421616]
23. Kent P, Mirkhil S, Keating J, Buchbinder R, Manniche C, Albert HB. The concurrent validity of brief screening questions for anxiety, depression, social isolation, catastrophization, and fear of movement in people with low back pain. Clin J Pain 2014 Jun;30(6):479-489. [doi: 10.1097/AJP.0000000000000010] [Medline: 24281277]
24. Philip L, Roberts A, Currie M, Mort A. Technology for older adults: maximising personal and social interaction: exploring opportunities for eHealth to support the older rural population with chronic pain. Scott Geogr J 2015 Jan 15;131(3-4):181-193. [doi: 10.1080/14702541.2014.978806]
25. Walrave M, Waeterloos C, Ponnet K. Reasons for nonuse, discontinuation of use, and acceptance of additional functionalities of a COVID-19 contact tracing app: cross-sectional survey study. JMIR Public Health Surveill 2022 Jan 14;8(1):e22113 [FREE Full text] [doi: 10.2196/22113] [Medline: 34794117]
26. Carroll JK, Moorhead A, Bond R, LeBlanc WG, Petrella RJ, Fiscella K. Who uses mobile phone health apps and does use matter? A secondary data analytics approach. J Med Internet Res 2017 Apr 19;19(4):e125 [FREE Full text] [doi: 10.2196/jmir.5604] [Medline: 28428170]
27. Lim D, Norman R, Robinson S. Consumer preference to utilise a mobile health app: A stated preference experiment. PLoS One 2020;15(2):e0229546 [FREE Full text] [doi: 10.1371/journal.pone.0229546] [Medline: 32084250]
28. Morsø L, Albert H, Kent P, Manniche C, Hill J. Translation and discriminative validation of the STarT Back Screening Tool into Danish. Eur Spine J 2011 Dec;20(12):2166-2173 [FREE Full text] [doi: 10.1007/s00586-011-1911-6] [Medline: 21769444]
29. Johnsen LG, Hellum C, Nygaard OP, Storheim K, Brox JI, Rossvoll I, et al. Comparison of the SF6D, the EQ5D, and the oswestry disability index in patients with chronic low back pain and degenerative disc disease. BMC Musculoskelet Disord 2013 Apr 26;14:148 [FREE Full text] [doi: 10.1186/1471-2474-14-148] [Medline: 23622053]
30. Faber JS, Al-Dhahir I, Reijnders T, Chavannes NH, Evers AWM, Kraal JJ, et al. Attitudes toward health, healthcare, and eHealth of people with a low socioeconomic status: a community-based participatory approach. Front Digit Health 2021;3:690182 [FREE Full text] [doi: 10.3389/fdgth.2021.690182] [Medline: 34713165]
31. Hitchman SC, Fong GT, Zanna MP, Thrasher JF, Chung-Hall J, Siahpush M. Socioeconomic status and smokers' number of smoking friends: findings from the International Tobacco Control (ITC) Four Country Survey. Drug Alcohol Depend 2014 Oct 01;143:158-166 [FREE Full text] [doi: 10.1016/j.drugalcdep.2014.07.019] [Medline: 25156228]

XSL•FO
RenderX

32. Tomioka K, Kurumatani N, Saeki K. The association between education and smoking prevalence, independent of occupation: a nationally representative survey in Japan. J Epidemiol 2020 Mar 05;30(3):136-142 [FREE Full text] [doi: 10.2188/jea.JE20180195] [Medline: 30828035]
33. Hansen A, Morsø L, Stochkendahl MJ, Andersen MKK, Schiøttz-Christensen B, Madsen SD, et al. Demographic and clinical characteristics of patients with low back pain in primary and secondary care settings in Southern Denmark. Scand J Prim Health Care 2023 Jun;41(2):152-159 [FREE Full text] [doi: 10.1080/02813432.2023.2196548] [Medline: 37154804]
34. Pan J, Dong H, Bryan-Kinns N. Perception and initial adoption of mobile health services of older adults in London: mixed methods investigation. JMIR Aging 2021 Nov 19;4(4):e30420 [FREE Full text] [doi: 10.2196/30420] [Medline: 34807836]

Abbreviations

LBP: low back pain
mHealth: mobile health
OR: odds ratio
PRO: patient-reported outcome
STROBE: Strengthening the Reporting of Observational Studies in Epidemiology

Edited by A Kushniruk; submitted 04.05.23; peer-reviewed by P Poulin, S Kheirinejad, C Baxter; comments to author 29.10.23; revised version received 12.12.23; accepted 20.01.24; published 12.02.24.

Please cite as:
Ipsen JA, Sandal LF, Chang NHS, Schiøttz-Christensen B, Søgaard K, Hansen A
Interest in mHealth Among Patients With Low Back Pain: Cross-Sectional Study
JMIR Hum Factors 2024;11:e48729
URL: https://humanfactors.jmir.org/2024/1/e48729
doi:10.2196/48729
PMID:38345837

©Jonas Ammundsen Ipsen, Louise Fleng Sandal, Natalie Hong Siu Chang, Berit Schiøttz-Christensen, Karen Søgaard, Anders Hansen. Originally published in JMIR Human Factors (https://humanfactors.jmir.org), 12.02.2024. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on https://humanfactors.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

A Mobile Health App to Support Home-Based Aerobic Exercise in Neuromuscular Diseases: Usability Study

Tim Veneman[1,2], MSc; Fieke Sophia Koopman[1,2], MD, PhD; Sander Oorschot[1,2], MSc; Pien G Koomen[1,2], PT; Frans Nollet[1,2], MD, PhD; Eric L Voorn[1,2], PhD

[1]Amsterdam University Medical Center location University of Amsterdam, Rehabilitation Medicine, Amsterdam, Netherlands

[2]Amsterdam Movement Sciences, Rehabilitation & Development, Amsterdam, Netherlands

Corresponding Author:
Eric L Voorn, PhD
Amsterdam University Medical Center location University of Amsterdam
Rehabilitation Medicine
Meibergdreef 9
Amsterdam, 1105 AZ
Netherlands
Phone: 31 20 56 64421
Email: e.l.voorn@amsterdamumc.nl

Abstract

Background: Home-based aerobic exercise in people with neuromuscular diseases (NMDs) has benefits compared to exercise in the hospital or a rehabilitation center because traveling is often cumbersome due to mobility limitations, and societal costs are lower. Barriers to home-based aerobic exercise include reduced possibilities for monitoring and lack of motivation. To overcome these and other barriers, we developed a mobile health app: Keep on training with ReVi (hereafter referred to as ReVi).

Objective: We aimed to determine the usability of the ReVi app.

Methods: Patients followed a 4-month, polarized, home-based aerobic exercise program on a cycle or rowing ergometer, with 2 low-intensity sessions and 1 high-intensity session per week supported by the ReVi app. The app collected training data, including heart rate and ratings of perceived exertion, provided real-time feedback on reaching target intensity zones, and enabled monitoring via an online dashboard. Physiotherapists instructed patients on how to use the ReVi app and supervised them during their training program. Patients and physiotherapists separately evaluated usability with self-developed questionnaires, including 9 questions on a 5-point Likert scale, covering the usability elements efficiency, effectiveness, and satisfaction.

Results: Twenty-nine ambulatory adult patients (n=19 women; mean age 50.4, SD 14.2 years) with 11 different slowly progressive NMDs participated. Both patients and physiotherapists (n=10) reported that the app, in terms of its efficiency, was easy to use and had a rapid learning curve. Sixteen patients (55%) experienced 1 or more technical issue(s) during the course of the exercise program. In the context of effectiveness, 23 patients (81%) indicated that the app motivated them to complete the program and that it helped them to exercise within the target intensity zones. Most patients (n=19, 70%) and physiotherapists (n=6, 60%) were satisfied with the use of the app. The median attendance rate was 88% (IQR 63%-98%), with 76% (IQR 69%-82%) of time spent within the target intensity zones. Four adverse events were reported, 3 of which were resolved without discontinuation of the exercise program.

Conclusions: The usability of the ReVi app was high, despite the technical issues that occurred. Further development of the app to resolve these issues is warranted before broader implementation into clinical practice.

(JMIR Hum Factors 2024;11:e49808) doi:10.2196/49808

KEYWORDS

neuromuscular disorders; endurance training; home-based exercise; eHealth; tele-rehabilitation; app; exercise; aerobic exercise; mhealth; mobile app; neuromuscular disease; usability

XSL•FO
RenderX

Introduction

Physical fitness is an important health marker [1,2] and is strongly associated with daily life functioning [3] and independent living [4] at older age. People with neuromuscular disease (NMD) often have reduced physical fitness caused not only by the underlying disease but also by an inactive lifestyle [5-7]. Aerobic exercise is an important aspect of rehabilitation treatment for NMD, as it contributes to improved physical fitness [8]. The integration of exercise programs into everyday life was recently identified as one of the major research priorities for individuals with NMD [9].

People with NMD usually perform their aerobic exercise program in a hospital, rehabilitation center, or physiotherapy practice under the direct supervision of a physiotherapist. However, center-based exercise may be cumbersome for individuals with NMD, who are often limited in their mobility. Moreover, center-based exercise requires the availability of physiotherapy staff, whose number is often limited as many countries are reducing health care services [10,11]. This amplifies the need for alternative modes of exercise intervention delivery that maintain high quality and effectiveness [12,13].

Transferring aerobic exercise from the hospital environment to the home or community may be a beneficial way to reduce travel time and societal costs. A recently developed training guide called B-FIT is an example of a home-based aerobic exercise program specifically developed for NMD [14]. Feasibility of the B-FIT exercise program has been demonstrated for different types of NMD, and patients and physiotherapists were satisfied with its use [14]. A barrier to use B-FIT was that some patients experienced the program as insufficiently challenging. This requires attention because poor motivation has been reported as a major barrier to exercise in people with NMD [7,15]. Furthermore, physiotherapists perceived initiation of the program as time-consuming; most of the worksheets, including exercise testing results and the training schedules, needed to be filled out by hand. A more general concern regarding exercise in the home environment is the reduced possibility to monitor exercise sessions. This is particularly important for the vulnerable population of people with NMD, as it may put them at risk for under- and overtraining.

To overcome these barriers to home-based aerobic exercise for people with NMD, we developed a mobile health (mHealth) app called Keep on training with ReVi (hereafter referred to as the ReVi app). The ReVi app aims to improve patients' adherence to the B-FIT exercise program by (1) offering a structured exercise program, (2) providing insight into training progression, and (3) improving motivation through auditory encouragement. For physiotherapists, the ReVi app aims to improve their opportunities for supervision by enabling them to monitor progress and provide feedback from a distance and also to reduce the time investment to initiate the exercise program.

The primary aim of this study was to assess the usability of the ReVi app for assisting and monitoring home-based aerobic exercise according to the B-FIT training guide in people with NMD. We also evaluated the attendance rate, the time spent within target intensity zones, and the occurrence of adverse effects.

Methods

Design

A multicenter prospective pilot study was conducted at the outpatient departments of rehabilitation medicine of 2 university hospitals and in 3 rehabilitation centers in the Netherlands. All centers were specialized in treatment of NMD. This study included 2 different cohorts; in one cohort, the ReVi app was applied as part of usual care at the Department of Rehabilitation Medicine at the Amsterdam UMC, location Amsterdam Medical Center (AMC). The other cohort consisted of patients using the ReVi app in the intervention group of an ongoing multicenter randomized controlled trial on the efficacy of a physical activity program, which combines the B-FIT aerobic exercise program and motivational interviewing coaching to improve physical fitness in people with NMD [16].

Ethical Considerations

The medical ethics review committee of the AMC waived the need for medical ethical approval for the usual care cohort, and approved the study protocol of the randomized controlled trial (NL62104.018.17). All patients provided informed consent.

Participants

The inclusion criteria applied to both cohorts were (1) diagnosis of a slowly progressive NMD, (2) age ≥18 years, and (3) possession of a smartphone or tablet. Exclusion criteria were (1) contraindication for being physically active, (2) inability to follow verbal or written instructions, and (3) insufficient competence in the Dutch language. In addition, patients in the randomized controlled trial had to be motivated to improve their reduced physical fitness and were excluded if they had participated in an exercise program for a period longer than 4 weeks in the past 6 months. For the purpose of this study, we included only data of patients who completed at least 12 of the 48 possible training sessions, to ensure sufficient experience with the use of the ReVi app to evaluate its usability. We aimed to include a total of 30 patients in this study.

Physiotherapists were included in the study if they supervised at least 1 patient. Physiotherapists that were already exposed to the B-FIT training guide followed a half-day training course to refresh their knowledge on the use of the B-FIT training program and learn the use of the ReVi app. Physiotherapists that were not exposed to the B-FIT training guide followed a full-day training course to learn both the B-FIT training program and the use of the app. Furthermore, they received an instruction manual with a step-by-step guide on the use of the app.

ReVi App

The ReVi app (Amsterdam UMC) was built by a company (everywhereIM BV) specialized in the development of medical apps. The app was available for iOS and Android and it was developed in the Dutch language. An expert group consisting of physiotherapists, rehabilitation physicians, exercise physiologists, patients with different types of NMD, and representatives of the Dutch Society of Muscle Diseases and of

XSL•FO
RenderX

the app builder actively participated in the development of the ReVi app. Expert group meetings were organized to discuss the aims of the app, to identify essential functionalities, and to provide feedback on so-called functional designs (on paper). The primary objective during this initial developmental phase was to create an app to assist a 16-week aerobic exercise regimen. If the study yields favorable results, the next developmental stage will be initiated to enhance the app's functionality and further explore the possibilities for offering longer-term support to home-based aerobic exercise. The data protection officer of Amsterdam UMC (location AMC) was also involved in the app's development process to ensure that personal data processing was organized in accordance with the General Data Protection Regulation (GDPR).

B-FIT Aerobic Exercise Program

The ReVi app was programmed with the B-FIT aerobic exercise program. This 16-week, polarized, home-based exercise program consisted of 2 low-intensity sessions below the anaerobic threshold (AT) and 1 high-intensity session above the AT per week. Patients visited the study center prior to the start, midway through, and after completion of the exercise program for a face-to-face meeting with their supervising physiotherapist. During each visit, an exercise test was executed. During the visits midway through and after completion of the exercise program, patients received feedback on training progress based on exercise testing results and based on data in the ReVi dashboard (see section App Description).

In the usual care cohort, target intensity zones were based on indirect assessment of the AT using ratings of perceived exertion (RPEs) during a submaximal exercise test [17]. In the randomized controlled trial cohort, target intensity zones were based on direct assessment of the AT during an exercise test through visual inspection of the gas exchange plots using the V-slope method [18]. If training based on heart rate was not feasible, for instance in patients using β-blocking agents, training was based on RPEs using the 6-20 Borg scale. Each training session consisted of several exercise intervals interspersed with recovery periods. Training sessions were performed in the home environment (eg, at home, in the gym, or at a physiotherapy practice) on a bicycle or rowing ergometer. A more detailed description of the B-FIT aerobic exercise program can be found in Multimedia Appendix 1 [14].

App Description

Physiotherapists created a personal account for a web-based dashboard that was used to create and manage ReVi app accounts of patients they supervised. The dashboard could be accessed using a desktop or laptop computer. Two-way verification using Google Authenticator (Google Inc) was required to sign in. The physiotherapists created patient accounts by sending a link to the patients' email addresses. Via this link, a password was created. Patients used the ReVi app on a mobile phone or tablet. Logging in to the ReVi app required their personal email address and password.

After signing in to the ReVi app, the home menu opened, from which 2 menus could be chosen: the Settings menu and the Training menu, which provided an overview of the program (Figure 1). Through the Settings menu the type of training could be chosen: training based on heart rate or based on Borg scale. For training based on heart rate, a Bluetooth connection with a heart rate monitor was established (in these cohorts, the device was the Polar H10; Polar Electro) and could be tested. Additionally, contact details of the physiotherapist were entered to enable patients to contact their therapists via the ReVi app.

The Training menu provided an overview of the training sessions (Figure 2). By selecting training sessions, the training protocol, including exercise intervals and recovery periods, was shown (Figure 3). During training sessions, the ReVi app guided users by illustrating their target intensity zones. In case of heart rate–based training, a heart rate chest strap was provided to the patient. The app was Bluetooth connected to the heart rate chest strap to continuously monitor heart rate (Figure 4). Patients rated their perceived exertion every final minute of the exercise interval or recovery period using the 6-20 Borg scale (Figure 5). During RPE-based training, patients rated their perceived exertion every minute. The ReVi app provided auditory feedback during training sessions. When patients trained within the target intensity zone, they were encouraged to continue. If the heart rate or Borg scale was not within the target intensity zone for at least 20 seconds, the ReVi app provided auditory instructions to increase or decrease the resistance. Directly after completion of the exercise session, an overview of the exercise results was shown (Figure 6).

Heart rate and Borg score data were saved by the ReVi app and sent to the web-based dashboard. Physiotherapists could access the training data of the patients they supervised; patients only had access to their own exercise data. The dashboard included, for each training session, a table with the percentage of time spent within the target intensity zones, the average heart rate for each exercise interval and recovery period, and the accompanying RPE (Figure 7). Additionally, a graph illustrated the actual heart rate or RPE with reference to the target intensity zones.

XSL•FO
RenderX

Figure 1. Screenshot of the ReVi app home screen.

Figure 2. Screenshot of the exercise program overview.

XSL•FO
RenderX

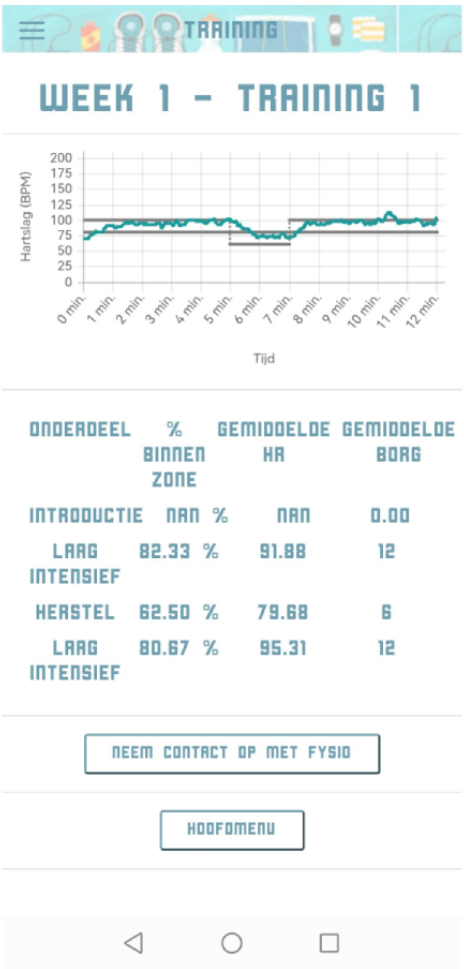
Figure 3. Screenshot of the exercise session protocol; this is an overview of the intensity and duration of each exercise or recovery bout.

Figure 4. Screenshot of the exercise session live screen; the actual achieved intensity (heart rate or rating of perceived exertion) and the target intensity zone during the exercise session are shown.

XSL•FO
RenderX

Figure 5. Screenshot of the Borg scale.

Figure 6. Screenshot of the exercise session results; the graph shows heart rate progression over time and the percentage of time spent within the target intensity zones.



XSL•FO
RenderX

Figure 7. Screenshot of the ReVi app web-based dashboard. The percentage of time spent within the target intensity zones is presented in the table. The graph shows the heart rate during the training session (solid line), as well as the target intensity zones (grey dashed lines). The ReVi app dashboard is only available in Dutch, but for this paper the screenshot was translated to English. HR: heart rate.

Segment	% within zone	Average HR	Zone	Borg
Warming-up	99	84	70 - 93	9
High intensive	76	123	122 - 133	12
Recovery	62	94	70 - 93	10
High intensive	73	123	122 - 133	13
Recovery	45	96	70 - 93	11
High intensive	83	125	122 - 133	13
Recovery	36	98	70 - 93	11

Outcomes

ReVi App Usability

The primary outcome was the usability of the ReVi app, defined according to the International Organization for Standardization (ISO) as follows: “Usability is the extent to which a product can be used by specified users to achieve specified goals with efficiency, effectiveness and satisfaction in a specified context of use” [19]. Efficiency refers to the resources expended in relation to the accuracy and completeness with which users achieve goals (eg, ease of use, learning time, and additional effort of using the ReVi app during training sessions). Effectiveness refers to the extent to which the ReVi app has completed its goals to motivate patients and support patients to train within the targeted heart rate zones. Satisfaction assesses positive or negative attitudes toward the use of the ReVi app [20].

Self-developed questionnaires were used to assess the usability of the ReVi app among patients and physiotherapists. The questionnaires were developed by the study team, which consisted of researchers, rehabilitation physicians, and a physiotherapist. The questionnaires were reviewed by 2 patients and another physiotherapist before the final version was developed. The questionnaires contained questions pertaining to the 3 major aspects of usability: efficiency, effectiveness, and satisfaction. The usability questionnaires for patients and physiotherapists included 12 and 13 questions, respectively, of which 2 were open questions (Multimedia Appendices 2 and 3). Nine of the closed questions were scored on a 5-point Likert scale (1=strongly disagree; 2=disagree; 3=neither agree nor disagree; 4=agree; 5=strongly agree). Patients filled in the questionnaire after their last completed training session; physiotherapists did so after completion by the last patient they supervised.

Attendance Rate and Time Within Target Intensity Zones

For assessing attendance rates and the time spent within target intensity zones, we used data collected in the ReVi app dashboard. The attendance rate was defined as the percentage of followed training sessions. From the followed training sessions, we determined the percentage of time spent within target intensity zones for low- and high-intensity exercise intervals combined and separately.

Adverse Events

Adverse events related to the exercise program, such as severe muscle fatigue, joint pain, or muscle pain, were recorded. Patients were instructed to contact the physiotherapist to report adverse events. In addition, physiotherapists checked for adverse events during each patient visit.

Data Analysis

Descriptive statistics are used to present patient and physiotherapist characteristics. The data from the questions that were scored on a 5-point Likert scale were reduced by combining “agree” and “strongly agree” responses to form an “agree” category, and response options of “strongly disagree”

XSL•FO
RenderX

and "disagree" were combined to form "disagree." Frequencies were calculated on the basis of the total number of responses to each question on the usability questionnaire and expressed as percentages. Data analysis was performed using SPSS (version 28.0; IBM Inc).

Results

Study Group

Between January 2020 and November 2021, 23 patients started their exercise program as part of the usual care cohort, of which 20 patients were included in the study. Three patients were excluded because they executed less than 12 exercise sessions. Reasons included technical problems with the ReVi app (n=1), medical issues (n=1), and a lack of motivation (n=1). Nine other patients participating in the ongoing randomized controlled trial were also included and started between July 2021 and December 2021.

Patient characteristics are shown in Table 1. Twenty-three patients were treated at the outpatient clinic of the Department of Rehabilitation of the Amsterdam UMC (location AMC), supervised by 6 physiotherapists. The other 6 patients were treated by 4 physiotherapists at Rehabilitation Center Klimmendaal (Arnhem; n=2), Basalt Rehabilitation Center (Leiden; n=2), University Medical Center Utrecht (Utrecht; n=1), and Sint Maartenskliniek (Nijmegen; n=1). Twenty-eight patients trained based on heart rate and 1 patient based on the Borg scale. Twenty-seven patients performed the exercise program using a bicycle ergometer and 2 patients used a rowing ergometer.

Table 1. Respondent profile.

Characteristics	Respondents
Patients (n=29)	
Age (years), mean (SD)	50.4 (14.2)
Female, n (%)	19 (66)
Sum score of manual muscle testing for the legs[a], median (range)	75 (60-80)
Peak workload baseline submaximal exercise test (watts), median (range)	100 (50-210)
Types of neuromuscular disorder, n	
Charcot-Marie-Tooth disease	7
Myotonic dystrophy	4
Nonspecific myopathy	4
Congenital dystrophy	3
Limb girdle muscular dystrophy	3
Mitochondrial myopathy	2
Inclusion body myositis	2
Becker muscular dystrophy	1
Postpolio syndrome	1
Dermatomyositis	1
Chronic inflammatory demyelinating polyradiculoneuropathy	1
Physiotherapists (n=10)	
Female, n (%)	7 (70)
Patients supervised in study (n), median (range)	2 (1-9)
Prior experience with use of the ReVi app, n (%)	4 (40)

[a]Sum score for muscle strength of the legs was calculated by adding 16 muscle groups. Each muscle group had a score between 0 and 5, and the sum score ranged from 0 to 80 [21].

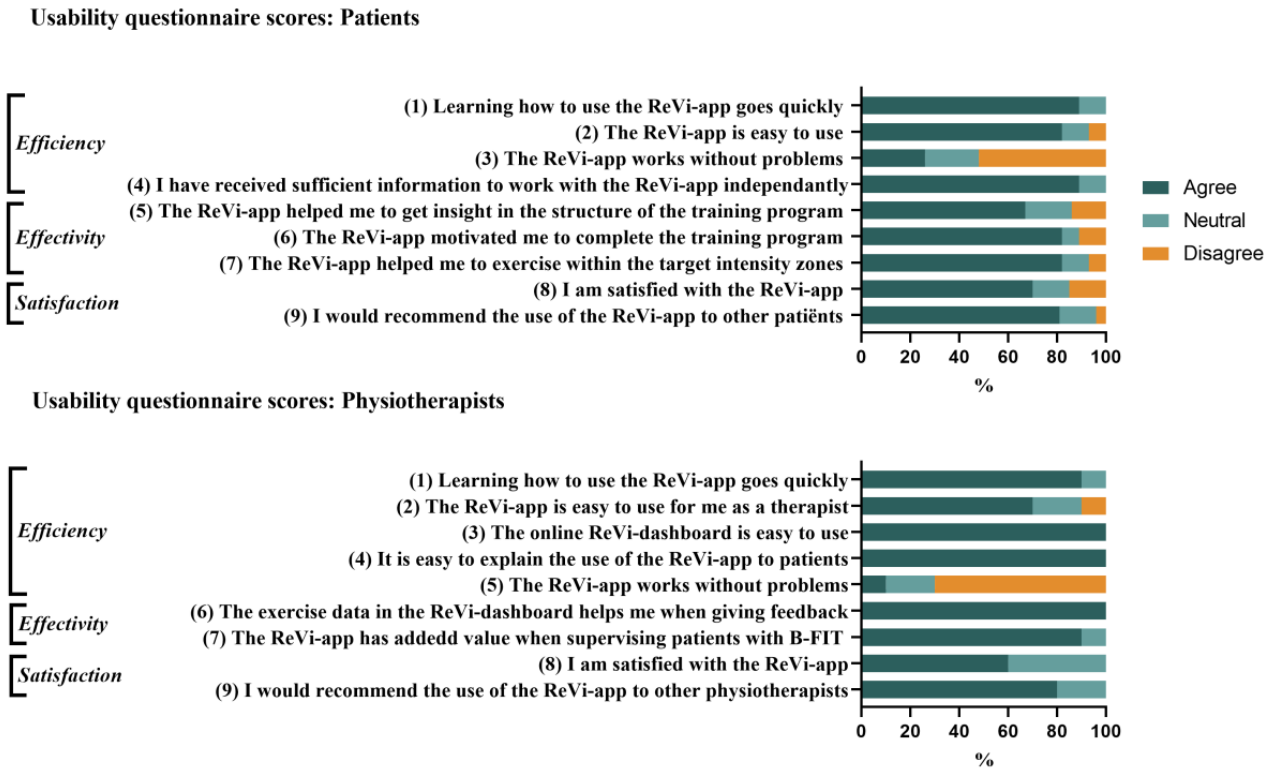
Primary Outcome

Usability

Twenty-seven patients and all 10 physiotherapists filled in and returned the usability questionnaire. Two patients did not return the usability questionnaire despite multiple requests. Questionnaire scores of patients and physiotherapists are presented in Figure 8.

XSL•FO
RenderX

Figure 8. Statements and outcomes for the patient usability questionnaire (n=27) and the physiotherapist usability questionnaire (n=10). Scores are given on a Likert scale, ranging from 1-5 (1 = strongly disagree, 5 = strongly agree). Frequency data were reduced by combining "agree" and " strongly agree" responses to form an "agree" category, and response options of "strongly disagree" and "disagree" were combined to form "disagree".



Efficiency

Twenty-four patients (89%) reported that learning how to use the ReVi app went quickly and 22 patients (81%) found that the ReVi app was easy to use. Seven patients (26%) agreed with the statement "the ReVi app works without problems." In 16 of the total of 29 patients (55%), 1 or more technical issues occurred during the course of the ReVi app training program. The most-reported technical issues were connection problems with the heart rate monitor and a bug in the app that hindered saving of exercise data in week 11 of the exercise program.

The majority of therapists reported that the ReVi app was easy to use (n=7, 70%) and all therapists found the use of the app easy to explain to patients. Nine therapists (90%) experienced technical issues using the ReVi app.

Effectiveness

Eighteen patients (67%) reported that the ReVi app provided insight into the structure of the exercise program. Twenty-two patients (81%) agreed that the app motivated them to complete the program and that it helped them to maintain exercise within the target intensity zones.

All therapists reported that the web-based dashboard helped them to provide feedback to patients and that the ReVi app had added value for supervision. The most important benefit reported by the physiotherapists was that the ReVi app allowed insight into the number of sessions that were followed and the exercise intensity that was achieved during training sessions.

Satisfaction

Nineteen patients (70%) were satisfied with the use of the ReVi app and 22 patients (81%) would recommend the use of the ReVi app to other patients with NMD. The most important reasons to recommend its use to others were that the app provided structure, helped them to train within the target intensity zones, and motivated them to complete their training sessions. The most-reported reason for patients not to recommend the ReVi app to others was the occurrence of technical issues.

Six therapists (60%) reported that they were satisfied with the use of the ReVi app and 8 therapists (80%) would recommend the use of the app to other physiotherapists. Reasons to recommend its use to others were that it was easy to use, enabled monitoring from a distance, and provided data that could be used to give tailored feedback to patients. The most-reported reason to not recommend the ReVi app to others was the technical issues that occasionally occurred when using the app.

Secondary Outcomes

Attendance and Time Within Target Intensity Zones

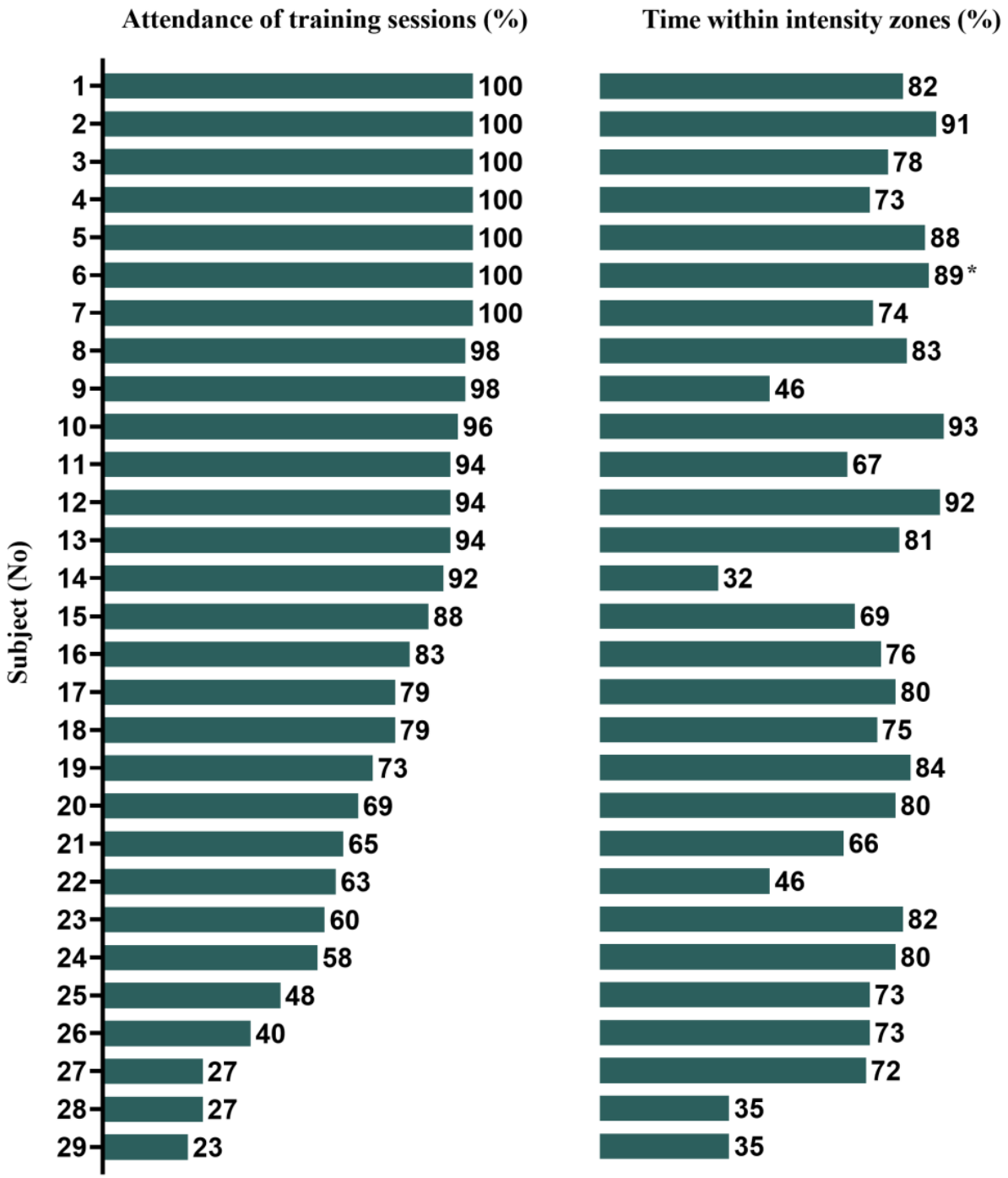
Twenty of the 29 patients (69%) completed the exercise program. Reasons for discontinuation among the other 9 patients were technical problems with the ReVi app (n=4), medical issues (n=2), closing of the local gym due to COVID-19 measures (n=2), and a lack of motivation (n=1).

Figure 9 shows the attendance rate for each patient, as well as the time spent within the target intensity zones. The median

XSL•FO
RenderX

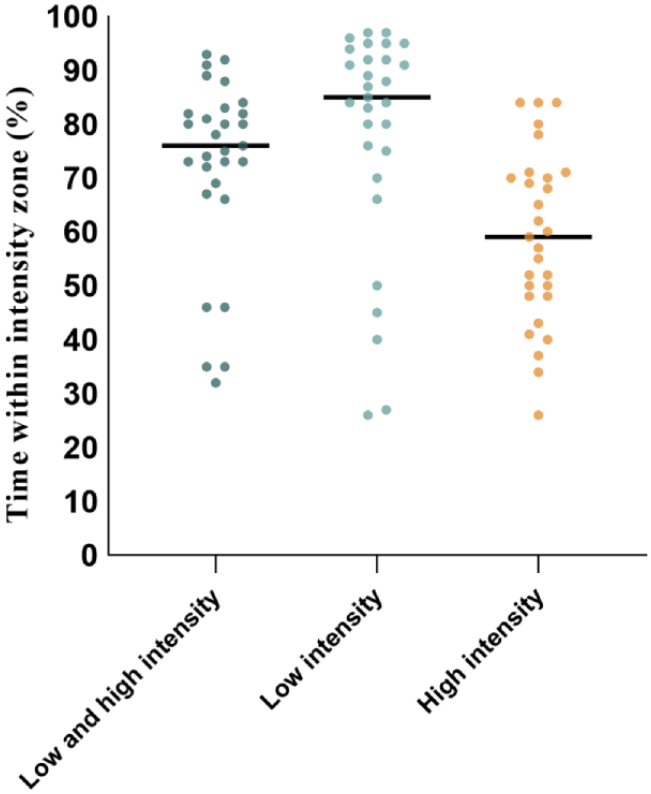
attendance rate was 88% (IQR 63%-98%). During the attended training sessions, patients spent a median of 76% (IQR 69%-82%) of the time within their target intensity zones (Figure 10). The median percentage of time spent within the low intensity zones was 85% (IQR 81%-92%), and in the high intensity zones it was 59% (IQR 45%-70%).

Figure 9. Attendance rates for individual patients ordered from most to least sessions, and the percentage of time spent within the target intensity zones during corresponding sessions. * patient trained based on Borg scale.



XSL•FO
RenderX

Figure 10. The percentage of time spent within the target intensity zones during A) both high and low intensity exercise intervals, B) low intensity exercise intervals and C) high intensity exercise intervals. Black lines indicate the median. Each dot represents a single patient.



Adverse Events

Four adverse events were reported: fatigue (n=2), knee joint pain (n=1), and high blood pressure during training (n=1). In the patient with high blood pressure during training, the rehabilitation physician and physiotherapist decided to terminate the exercise program. The other 3 adverse events were resolved without discontinuation of the exercise program.

Discussion

Principal Findings

This study provides insight into the usability of the ReVi app among people with NMD to support home-based aerobic exercise according to the B-FIT training program. The different components of usability, including efficiency, effectiveness, and satisfaction, were all judged as good by physiotherapists and patients, despite the occurrence of technical issues.

Patients were generally positive about the efficiency of the ReVi app due to its rapid learning curve and ease of use. Patients could independently work with the app based on the instructions that they received from their treating physiotherapist. Adequate instructions are known to be a key facilitator of patient engagement with mHealth apps [22]. With regards to its effectiveness, patients reported that the most important goals of the ReVi app were achieved: its use motivated them to complete the exercise program and helped them to exercise within their target intensity zones. These outcomes were supported by the findings that patients attended the majority of training sessions and spent most time within the target intensity zones. Patients were mostly satisfied with the use of the app, which concurs with other studies on apps supporting home-based physical exercise programs in amyotrophic lateral sclerosis, which is a rapidly progressive type of NMD [23], and a variety of other patient populations [24-26].

Physiotherapists were positive about the efficiency of the ReVi app. This was mainly due to the rapid learning curve and its ease of use; a half- or full-day training course was required for physiotherapists to learn how to work with the app and the B-FIT training guide, depending on prior experience with B-FIT. In terms of effectiveness, physiotherapists reported that the most important goals of the ReVi app were achieved. They found the app helpful when monitoring patients during their home-based program, mainly because it enabled them to provide feedback based on exercise data. They were generally satisfied with the use of the ReVi app and would recommend the use of the app to other physiotherapists.

While efficiency, effectiveness, and user satisfaction were overall judged as positive, one of the efficiency items was clearly judged as insufficient: 55% (n=16) of the patients and 90% (n=9) of the physiotherapists experienced technical issues. Most of these issues were solved, but in all cases, this required the help of a physiotherapist, researcher, or software developer. Technical issues are known to negatively impact usability and decrease adherence and engagement with mHealth tools [22]. They often cause patients to stop their mHealth interventions, leading to high dropout rates, and they are reported as a main barrier to further implementation of mHealth or eHealth apps [22,27]. This is consistent with our finding that the most important reason for discontinuation of the exercise program was when the ReVi app did not function well. Therefore, resolving technical issues is an important concern for further

XSL•FO
RenderX

implementation of the ReVi app in clinical rehabilitation practice on a broader scale. This also underlines the importance of offering technical support when using mHealth tools such as the ReVi app [28]. Physiotherapists could play an important role in this, but that would require sufficient proficiency with mHealth. Moreover, considering the limited availability of physiotherapy personnel, it is essential for successful implementation of mHealth tools like the ReVi app to minimize technical issues and provide access to additional technical support for more complex problems.

The attendance rate, time within target intensity zones, and adverse events found in this study suggest that training in the home environment with the help of the ReVi app is a good alternative to center-based training. The attendance rate of 88% and time within target intensity zones of 76% are in line with adherence rates found in other studies evaluating aerobic exercise programs for NMD [29-37] that were mostly conducted in a hospital or rehabilitation center. Comparison of the attendance rate and time within target intensity zones between this study and past studies on exercise for NMD is hampered by incomplete or absent descriptions of adherence assessment methods in most other studies. In some studies, it is unclear if reported values are for attendance rates, the time spent within exercise zones, or training time. Moreover, some studies excluded patients who dropped out, leading to overestimated adherence. In this light, the attendance rate in our study may have been impacted by excluding patients who performed less than 12 exercise sessions and by the finding that some patients performed several training sessions without using the ReVi app. Despite these uncertainties, the attendance rate and time within target intensity zones found in our study seem to be in line with values reported in other aerobic exercise studies. The limited number of adverse events reported in this study also concurs with other studies on center-based aerobic exercise programs for NMD [32,38,39]. This further strengthens the notion that home-based aerobic exercise supported by the ReVi app may be considered a safe and feasible alternative for center-based exercise programs, which is in line with earlier research in telemonitoring of home-based exercise for amyotrophic lateral sclerosis [23].

Limitations

Patients with a positive attitude toward the use of mHealth may have been more inclined to participate in this study, causing selection bias and limiting generalizability to people with NMD and less affinity for mHealth. Also, most patients trained under supervision of a physiotherapist experienced in treating patients with NMD, which limits generalizability of our results to other health care settings, such as primary care physiotherapy practices.

Future Studies

Implementation of mHealth, such as with the ReVi app, in rehabilitation care presents some major challenges, such as the comfort of patients and therapists with the use of technology, legal and ethical considerations regarding patient monitoring and the protection of privacy rights, and integration of mHealth tools into current working protocols [40,41]. Additionally, specific application design requirements have to be considered for NMD patients who experience reduced hand functionality due to muscle weakness. These requirements may include sufficiently large buttons and input fields. As a consequence of these challenges, the scientific literature on telerehabilitation in NMD patients is still limited [42,43]. To enable the broader implementation of mHealth in clinical practice, research is warranted into other facilitators of and barriers to the implementation of mHealth specific to neuromuscular rehabilitation.

Conclusions

The usability of the ReVi app in terms of perceived efficiency, effectiveness, and user satisfaction is high, despite the occurrence of technical issues. Combined with the high attendance rate and time spent within target intensity zones and low number of adverse events, the ReVi app can be considered a promising tool to support home-based aerobic exercise in rehabilitation practice for NMD. Further development of the ReVi app to resolve technical issues is warranted before broader implementation into clinical rehabilitation practice.

Acknowledgments

We thank all participating patients and physiotherapists. This research project was funded by IPSEN Pharmaceutics BV.

Conflicts of Interest

None declared.

Multimedia Appendix 1
B-FIT exercise program.
[DOCX File , 110 KB - humanfactors_v11i1e49808_app1.docx]

Multimedia Appendix 2
Patient usability questionnaire.
[DOCX File , 36 KB - humanfactors_v11i1e49808_app2.docx]

XSL•FO
RenderX

Multimedia Appendix 3
Therapist usability questionnaire.
[DOCX File , 35 KB - humanfactors_v11i1e49808_app3.docx]

References

1. Eriksen L, Curtis T, Grønbæk M, Helge JW, Tolstrup JS. The association between physical activity, cardiorespiratory fitness and self-rated health. Prev Med 2013 Dec;57(6):900-902. [doi: 10.1016/j.ypmed.2013.09.024] [Medline: 24099878]
2. Ortega FB, Ruiz JR, Castillo MJ, Sjöström M. Physical fitness in childhood and adolescence: a powerful marker of health. Int J Obes (Lond) 2008 Jan;32(1):1-11. [doi: 10.1038/sj.ijo.0803774] [Medline: 18043605]
3. Arnett SW, Laity JH, Agrawal SK, Cress ME. Aerobic reserve and physical functional performance in older adults. Age Ageing 2008 Jul;37(4):384-389. [doi: 10.1093/ageing/afn022] [Medline: 18287178]
4. Cress ME, Meyer M. Maximal voluntary and functional performance levels needed for independence in adults aged 65 to 97 years. Phys Ther 2003 Jan;83(1):37-48. [Medline: 12495411]
5. Aitkens S, Kilmer DD, Wright NC, McCrory MA. Metabolic syndrome in neuromuscular disease. Arch Phys Med Rehabil 2005 May;86(5):1030-1036. [doi: 10.1016/j.apmr.2004.09.012] [Medline: 15895353]
6. McDonald CM. Physical activity, health impairments, and disability in neuromuscular disease. Am J Phys Med Rehabil 2002 Nov;81(11 Suppl):S108-S120. [doi: 10.1097/00002060-200211001-00012] [Medline: 12409816]
7. Phillips M, Flemming N, Tsintzas K. An exploratory study of physical activity and perceived barriers to exercise in ambulant people with neuromuscular disease compared with unaffected controls. Clin Rehabil 2009 Aug;23(8):746-755. [doi: 10.1177/0269215509334838] [Medline: 19506002]
8. Cup EH, Pieterse AJ, Ten Broek-Pastoor JM, Munneke M, van Engelen BG, Hendricks HT, et al. Exercise therapy and other types of physical therapy for patients with neuromuscular diseases: a systematic review. Arch Phys Med Rehabil 2007 Nov;88(11):1452-1464. [doi: 10.1016/j.apmr.2007.07.024] [Medline: 17964887]
9. Nierse CJ, Abma TA, Horemans AMC, van Engelen BGM. Research priorities of patients with neuromuscular disease. Disabil Rehabil 2013 Mar;35(5):405-412. [doi: 10.3109/09638288.2012.694964] [Medline: 22747255]
10. Ratcliffe J, Laver K, Couzner L, Cameron ID, Gray L, Crotty M. Not just about costs: the role of health economics in facilitating decision making in aged care. Age Ageing 2010 Jul;39(4):426-429. [doi: 10.1093/ageing/afq041] [Medline: 20444804]
11. World report on ageing and health. World Health Organization. 2015. URL: https://www.who.int/publications/i/item/9789241565042 [accessed 2024-02-27]
12. Lai B, Young HJ, Bickel CS, Motl RW, Rimmer JH. Current trends in exercise intervention research, technology, and behavioral change strategies for people with disabilities: A scoping review. Am J Phys Med Rehabil 2017 Oct;96(10):748-761. [doi: 10.1097/PHM.0000000000000743] [Medline: 28398967]
13. Lai B, Kim Y, Wilroy J, Bickel CS, Rimmer JH, Motl RW. Sustainability of exercise intervention outcomes among people with disabilities: a secondary review. Disabil Rehabil 2019 Jun;41(13):1584-1595. [doi: 10.1080/09638288.2018.1432704] [Medline: 29409367]
14. Voorn EL, Koopman FS, Nollet F, Brehm MA. Individualized aerobic exercise in neuromuscular diseases: a pilot study on the feasibility and preliminary effectiveness to improve physical fitness. Phys Ther 2021 Mar 03;101(3):pzaa213 [FREE Full text] [doi: 10.1093/ptj/pzaa213] [Medline: 33332538]
15. Essery R, Geraghty AWA, Kirby S, Yardley L. Predictors of adherence to home-based physical therapies: a systematic review. Disabil Rehabil 2017 Mar;39(6):519-534. [doi: 10.3109/09638288.2016.1153160] [Medline: 27097761]
16. Oorschot S, Brehm MA, van Groenestijn AC, Koopman FS, Verhamme C, Eftimov F, et al. Efficacy of a physical activity programme combining individualized aerobic exercise and coaching to improve physical fitness in neuromuscular diseases (I'M FINE): study protocol of a randomized controlled trial. BMC Neurol 2020 May 13;20(1):184 [FREE Full text] [doi: 10.1186/s12883-020-01725-0] [Medline: 32404133]
17. Voorn EL, Gerrits KH, Koopman FS, Nollet F, Beelen A. Determining the anaerobic threshold in postpolio syndrome: comparison with current guidelines for training intensity prescription. Arch Phys Med Rehabil 2014 May;95(5):935-940. [doi: 10.1016/j.apmr.2014.01.015] [Medline: 24491465]
18. Wasserman K, Hansen J, Sue D, Stringer W, Whipp B. Principles of Exercise Testing and Interpretation, Including Pathophysiology and Clinical Applications. Philadelphia, PA: Lippincott Williams & Wilkins; 2005.
19. Ergonomic requirements for office work with visual display terminals (VDTs) - Visual display requirements. British Standards Institution. 1998. URL: https://knowledge.bsigroup.com/products/ergonomic-requirements-for-office-work-with-visual-display-terminals-vdts-visual-display-requirements?version=standard&tab=preview [accessed 2024-02-05]
20. Harrison R, Flood D, Duce D. Usability of mobile applications: literature review and rationale for a new usability model. J Interact Sci 2013;1(1):1 [FREE Full text] [doi: 10.1186/2194-0827-1-1]
21. Secer H, Daneyemez M, Gonul E, Izci Y. Medical research council: aids to examination of the peripheral nervous system. UK Research and Innovation. 1976. URL: https://www.ukri.org/wp-content/uploads/2021/12/MRC-011221-AidsToTheExaminationOfThePeripheralNervousSystem.pdf [accessed 2024-02-27]

XSL•FO
RenderX

22. Simblett S, Greer B, Matcham F, Curtis H, Polhemus A, Ferrão J, et al. Barriers to and facilitators of engagement with remote measurement technology for managing health: systematic review and content analysis of findings. J Med Internet Res 2018 Jul 12;20(7):e10480 [FREE Full text] [doi: 10.2196/10480] [Medline: 30001997]
23. Braga AC, Pinto A, Pinto S, de Carvalho M. Tele-monitoring of a home-based exercise program in amyotrophic lateral sclerosis: a feasibility study. Eur J Phys Rehabil Med 2018 Jun;54(3):501-503 [FREE Full text] [doi: 10.23736/S1973-9087.18.05129-8] [Medline: 29363923]
24. Arensman R, Kloek C, Pisters M, Koppenaal T, Ostelo R, Veenhof C. Patient perspectives on using a smartphone app to support home-based exercise during physical therapy treatment: qualitative study. JMIR Hum Factors 2022 Sep 13;9(3):e35316 [FREE Full text] [doi: 10.2196/35316] [Medline: 36098993]
25. Bizzarini E, Chittaro L, Frezza M, Polo M, Malisan C, Menosso R, et al. A mobile app for home-based exercise in spinal cord injured persons: Proposal and pilot study. Digit Health 2022;8:20552076211070724 [FREE Full text] [doi: 10.1177/20552076211070724] [Medline: 35140978]
26. Mahmood A, Blaizy V, Verma A, Stephen Sequeira J, Saha D, Ramachandran S, et al. Acceptability and attitude towards a mobile-based home exercise program among stroke survivors and caregivers: a cross-sectional study. Int J Telemed Appl 2019;2019:5903106 [FREE Full text] [doi: 10.1155/2019/5903106] [Medline: 31186627]
27. Hardiker NR, Grant MJ. Factors that influence public engagement with eHealth: A literature review. Int J Med Inform 2011 Jan;80(1):1-12 [FREE Full text] [doi: 10.1016/j.ijmedinf.2010.10.017] [Medline: 21112244]
28. Talboom-Kamp EP, Verdijk NA, Kasteleyn MJ, Numans ME, Chavannes NH. From chronic disease management to person-centered eHealth; a review on the necessity for blended care. Clin eHealth 2018 Mar;1(1):3-7 [FREE Full text] [doi: 10.1016/j.ceh.2018.01.001]
29. Koopman FS, Voorn EL, Beelen A, Bleijenberg G, de Visser M, Brehm MA, et al. No reduction of severe fatigue in patients with postpolio syndrome by exercise therapy or cognitive behavioral therapy: results of an RCT. Neurorehabil Neural Repair 2016 Jun;30(5):402-410. [doi: 10.1177/1545968315600271] [Medline: 26253175]
30. Voet N, Bleijenberg G, Hendriks J, de Groot I, Padberg G, van Engelen B, et al. Both aerobic exercise and cognitive-behavioral therapy reduce chronic fatigue in FSHD: an RCT. Neurology 2014 Nov 18;83(21):1914-1922. [doi: 10.1212/WNL.0000000000001008] [Medline: 25339206]
31. Sveen ML, Jeppesen TD, Hauerslev S, Køber L, Krag TO, Vissing J. Endurance training improves fitness and strength in patients with Becker muscular dystrophy. Brain 2008 Nov;131(Pt 11):2824-2831. [doi: 10.1093/brain/awn189] [Medline: 18776212]
32. Ernstoff B, Wetterqvist H, Kvist H, Grimby G. Endurance training effect on individuals with postpoliomyelitis. Arch Phys Med Rehabil 1996 Sep;77(9):843-848. [doi: 10.1016/s0003-9993(96)90268-3] [Medline: 8822672]
33. El Mhandi L, Millet GY, Calmels P, Richard A, Oullion R, Gautheron V, et al. Benefits of interval-training on fatigue and functional capacities in Charcot-Marie-Tooth disease. Muscle Nerve 2008 May;37(5):601-610. [doi: 10.1002/mus.20959] [Medline: 18335470]
34. Andersen G, Heje K, Buch AE, Vissing J. High-intensity interval training in facioscapulohumeral muscular dystrophy type 1: a randomized clinical trial. J Neurol 2017 Jun;264(6):1099-1106. [doi: 10.1007/s00415-017-8497-9] [Medline: 28470591]
35. Andersen G, Prahm KP, Dahlqvist JR, Citirak G, Vissing J. Aerobic training and postexercise protein in facioscapulohumeral muscular dystrophy: RCT study. Neurology 2015 Aug 04;85(5):396-403. [doi: 10.1212/WNL.0000000000001808] [Medline: 26156512]
36. Orngreen MC, Olsen DB, Vissing J. Aerobic training in patients with myotonic dystrophy type 1. Ann Neurol 2005 May;57(5):754-757. [doi: 10.1002/ana.20460] [Medline: 15852373]
37. Sveen M, Jeppesen TD, Hauerslev S, Krag TO, Vissing J. Endurance training: an effective and safe treatment for patients with LGMD2I. Neurology 2007 Jan 02;68(1):59-61. [doi: 10.1212/01.wnl.0000250358.32199.24] [Medline: 17200494]
38. Alemo Munters L, Dastmalchi M, Andgren V, Emilson C, Bergegård J, Regardt M, et al. Improvement in health and possible reduction in disease activity using endurance exercise in patients with established polymyositis and dermatomyositis: a multicenter randomized controlled trial with a 1-year open extension followup. Arthritis Care Res (Hoboken) 2013 Dec;65(12):1959-1968 [FREE Full text] [doi: 10.1002/acr.22068] [Medline: 23861241]
39. Cejudo P, Bautista J, Montemayor T, Villagómez R, Jiménez L, Ortega F, et al. Exercise training in mitochondrial myopathy: a randomized controlled trial. Muscle Nerve 2005 Sep;32(3):342-350. [doi: 10.1002/mus.20368] [Medline: 15962332]
40. Galetsi P, Katsaliaki K, Kumar S. Exploring benefits and ethical challenges in the rise of mHealth (mobile healthcare) technology for the common good: An analysis of mobile applications for health specialists. Technovation 2023 Mar;121:102598 [FREE Full text] [doi: 10.1016/j.technovation.2022.102598]
41. DeRuyter F, Jones M, Morris J. Mobile Health and mobile rehabilitation for people with disabilities: current state, challenges and opportunities. In: Computers Helping People with Special Needs: 17th International Conference. 2020 Presented at: ICCHP 2020; September 9-11, 2020; Lecco, Italy p. 17. [doi: 10.1007/978-3-030-58805-2_26]
42. Giannotta M, Petrelli C, Pini A. Telemedicine applied to neuromuscular disorders: focus on the COVID-19 pandemic era. Acta Myol 2022 Mar;41(1):30-36 [FREE Full text] [doi: 10.36185/2532-1900-066] [Medline: 35465343]

XSL•FO
RenderX

43. Costamagna G, Abati E, Bresolin N, Comi GP, Corti S. Management of patients with neuromuscular disorders at the time of the SARS-CoV-2 pandemic. J Neurol 2021 May;268(5):1580-1591 [FREE Full text] [doi: 10.1007/s00415-020-10149-2] [Medline: 32804279]

Abbreviations

AMC: Amsterdam Medical Center
AT: anaerobic threshold
GDPR: General Data Protection Regulation
ISO: International Organization for Standardization
mHealth: mobile health
NMD: neuromuscular disease
RPE: rating of perceived exertion

Edited by A Kushniruk; submitted 12.06.23; peer-reviewed by Y Chu, G Ramdharry; comments to author 31.08.23; revised version received 21.09.23; accepted 20.01.24; published 15.03.24.

Please cite as:
Veneman T, Koopman FS, Oorschot S, Koomen PG, Nollet F, Voorn EL
A Mobile Health App to Support Home-Based Aerobic Exercise in Neuromuscular Diseases: Usability Study
JMIR Hum Factors 2024;11:e49808
URL: https://humanfactors.jmir.org/2024/1/e49808
doi:10.2196/49808
PMID:38488838

©Tim Veneman, Fieke Sophia Koopman, Sander Oorschot, Pien G Koomen, Frans Nollet, Eric L Voorn. Originally published in JMIR Human Factors (https://humanfactors.jmir.org), 15.03.2024. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on https://humanfactors.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Short Paper

Designing a Novel Digitally Delivered Antiracism Intervention for Mental Health Clinicians: Exploratory Analysis of Acceptability

Tashalee Rushell Brown[1], MD, PhD; Habiba Amir[1]; Drew Hirsch[1]; Madeline Owens Jansen[1], MD, MPH

Department of Psychiatry and Biobehavioral Sciences, University of California, Los Angeles, Los Angeles, CA, United States

Corresponding Author:
Tashalee Rushell Brown, MD, PhD
Department of Psychiatry and Biobehavioral Sciences
University of California, Los Angeles
760 Westwood Plaza, 48-257
Los Angeles, CA, 90024
United States
Phone: 1 310 206 0188
Email: tashaleebrown@mednet.ucla.edu

Abstract

Background: There is a great need for evidence-based antiracism interventions targeting mental health clinicians to help mitigate mental health disparities in racially and ethnically minoritized groups.

Objective: This study provides an exploratory analysis of mental health clinicians' perspectives on the acceptability of a web-based antiracism intervention.

Methods: Mental health clinicians were recruited from a single academic medical center through outreach emails. Data were collected through individual 30-minute semistructured remote video interviews with participants, then recorded, transcribed, and analyzed using content analysis.

Results: A total of 12 mental health clinicians completed the study; 10 out of 12 (83%) were female candidates. Over half (7/12, 58%) of the respondents desired more robust antiracism training in mental health care. Regarding the web-based antiracism intervention, (8/12, 67%) enjoyed the digitally delivered demo module, (7/12, 58%) of respondents suggested web-based content would be further enhanced with the addition of in-person or online group components.

Conclusions: Our results suggest a strong need for additional antiracist training for mental health clinicians. Overall, participants responded favorably to novel web-based delivery methods for an antiracism intervention. These findings provide important support for future development and pilot testing of a large-scale digitally enhanced antiracist curriculum targeting mental health clinicians.

(JMIR Hum Factors 2024;11:e52561) doi:10.2196/52561

KEYWORDS

acceptability; antiracism; clinicians; intervention; interview study; mental health; psychiatry residents; racism; social workers; web-based technology

Introduction

Racism or expressions of discrimination are often rooted in implicit bias and stigmatizing beliefs [1]. Currently, racism is known to be a key driver of mental health inequities in ethnoracially minoritized groups who may be victims of discrimination [2]. Such experiences often lead to negative mental health outcomes [2]. Current evidence suggests that Black, Indigenous, and people of color youth and adults experience highly disproportionate rates of delayed diagnosis and treatment of autism spectrum disorder, overdiagnosis of conduct disorder, and underdiagnosis of attention-deficit/hyperactivity disorder [3], overdiagnosis of schizophrenia, overuse of antipsychotics with long-term medical consequences, and the underdiagnosis and treatment of depression [4]. Antiracism is the practice of actively opposing the effects of racism through institutional policies and individual behaviors [5]. Several recent systematic and scoping reviews on antiracism interventions in mental health professions have identified only one relevant randomized pilot study to date [5,6]. Of additional importance is that the authors found significant variability in training methodology, variability of intervention duration, and a lack of sufficient efficacy measurements to

XSL•FO
RenderX

evaluate existent antiracism interventions [5,6]. Thus, despite the strong need for evidence-based antiracism interventions targeting racial bias among mental health clinicians, such interventions remain underdeveloped and understudied in the literature. Within this context, evidence-based strategies, such as those based in cognitive-behavioral frameworks, have shown promise in addressing prejudiced thoughts, feelings, and behaviors but have yet to be applied to clinicians [4]. Notably, the delivery of tailored psychoeducational content such as this, has the potential to be greatly enhanced by digital design and delivery methods [3,7]. This is especially poignant given that web-based technologies are known to further augment interventional implementation structures with regard to both flexibility and sustainability [3,7].

Against the backdrop of a profound dearth of evidence-based antiracism interventions targeting mental health professionals, this study aims to explore aspects of the acceptability of a novel digitally delivered intervention of this sort [7]. Grounded in a strongly evidenced implementation science framework and through a dynamic and iterative process of evaluation, we explored facets of intervention acceptability regarding content, delivery, and implementation strategies [8]. Semistructured interviews were designed to elicit additional perceptions and attitudes among mental health professionals regarding gaps and opportunities in their current training on antiracism. Findings have the potential to be incorporated into future modifications of the intervention in order to optimize the feasibility and acceptability of large-scale randomized control pilot trials.

Methods

Overview

Participants were residents, fellows, and social workers specializing in mental health care. They were recruited from a single academic medical center in California through a remote method, which included outreach emails. Written, informed consent was obtained from all participants. Participants were compensated through US $50 gift cards and water bottles. Data were collected through individual 30-minute semistructured remote video interviews with participants, which were recorded and transcribed for analysis. Semistructured interview questions were developed based on the clinical experience and literature review conducted by MOJ and TRB (Table 1).

The semistructured interview featured a presentation of a digital demo module of the cognitive behavioral therapy (CBT)–based intervention, which discussed core beliefs that may be harmful in the treatment of patients with mental health conditions. The module features real-world examples, teaches a key concept of intervention, presents examples of self-monitoring, and provides a visual outline of the engagement and reward components. Data were qualitatively analyzed using inductive coding and thematic analysis methods [9] using the Atlas.ti (Scientific Software Development GmbH) software by 2 independent coders (HA and DH). Identified codes and themes were reviewed and consolidated by the leading authors (MOJ and TRB) until consensus was achieved. The number of respondents mentioning each code or theme was reported.

Table 1. Semistructured interview questions and probes.

Domain	Questions	Follow-up probes
Sociodemographic information	• What is your profession? • Where are you in your training?	• How would you describe your race and gender?
Definitions and thoughts of antiracism	• How would you define antiracism? • What terms do you like to use to describe or discuss racism and antiracism?	• Are you comfortable talking with coworkers or supervisors about racism and antiracism?
Strengths of the current antiracism training	• What are the strengths of your medical training thus far with regard to antiracism?	• What training, educational tools or courses have you benefited from in medical school or at the postgraduate level?
Weaknesses of the current antiracism training	• What are the weaknesses of your medical training thus far with regard to antiracism?	• What additional support, educational tools, or resources would help elevate your clinical skills to provide equitable care to diverse populations?

Feedback on the demo module	• How would you describe your experience going through the demo module?
	• What format would you prefer for antiracism training (in person, online, zoom)?
	• Would you have 10-15 minutes to dedicate to this specific type of antiracism learning?
	• Would seeing a report of potential bias in your electronic health record make you more or less likely to complete antiracism training? Why or why not?
		• What did you find helpful about the demo module?
		• What would make it more helpful?
		• How would you feel about your organization using digital means (online) in the form of self-directed modules to provide antiracism training?
		• Would 10-15 mins, once per week, for 6 weeks seem manageable?
		• A report of potential bias may include:
		  • Frequency of biased statements in notes
		  • Racial disparities in prescribing patterns

XSL•FO
RenderX

Ethical Considerations

This study was approved by the University of California, Los Angeles Institutional Review Board (IRB#22-001632-AM-00002).

Results

A total of 12 mental health clinicians (psychiatry residents, fellows, and social workers) completed the semistructured interviews. The participant characteristics included: female candidates (10/12, 83%), male candidates (2/12, 17%), and Asian (5/12, 42%), Black (2/12, 17%), Hispanic or Latinx (1/12, 8%), Middle Eastern (1/12, 8%), multiracial (1/12, 8%), White (1/12, 8%), and other (1/12, 8%) candidates.

The results of the content and thematic analysis are summarized in Table 2, but major themes are highlighted as follows: the majority of participants (7/12, 58%) desired more robust antiracism training in mental health care. With regard to the demo module, the majority (8/12, 67%) enjoyed the module, (6/12, 50%) found it to be well-organized, and (11/12, 92%) felt the time commitment to be manageable. Many participants particularly enjoyed the CBT-based content (4/12, 33%), especially the daily self-reflection log (4/12, 33%). About 4 participants expressed a preference for an online self-directed structure, and 7/12 (58%) participants suggested that online content could be enhanced with an in-person or group component. Lastly, 4 participants communicated ways to improve participant engagement through the digital modality, including offering incentives, sharing personal experiences, and recording progress.

XSL•FO
RenderX

Table 2. Themes and representative quotes from semistructured interviews.

Question	Themes	Quotes
Discussions about race	• Comfortable (5 mentions) • Somewhat comfortable (2 mentions) • Not very comfortable (4 mentions)	• I think so … I have to admit that oftentimes in the face of authority figures, it can be challenging…, it can get tiring though, when you're one of the few faces of color, or if you're like, the only Black person in the room…
Definitions of racism and antiracism	• Active advocacy against racism (11 mentions) • Racism as all-encompassing and systemic (5 mentions) • Self-awareness of antiracism (4 mentions)	• Antiracism, specifically, is a life-long journey, being aware of racial dynamics and disparities and power dynamics, I see it as, like, a modifiable factor.
Previous antiracism training	• Beneficial, in-depth discussions and courses at some point (11 mentions) • Limitations of training format and practicality (8 mentions) • Strong training in residency (7 mentions) • Minimal or no antiracism training (6 mentions) • Insufficient institutional support (5) • Beneficial scenario training (4 mentions)	• A lot of it is very theoretical; less of it is practical in the sense of, you know, in a specific situation. • I feel like there's a lot of, like, resident-driven antiracism efforts … justice, equity, diversity, and inclusion groups… • Anti-racist work has been performative, ...there was too high a burden on faculty of color…
Antiracism training needs	• More robust training and resources (7 mentions) • Integration of representation and lived experiences (4 mentions) • Accessible language (4 mentions) • Integrate translational social sciences in curriculum (3 mentions) • Increase cultural competency (2 mentions) • Mitigate minority tax (2 mentions)	• Just hiring, you know, more faculty of color, I feel that the best ways I've learned have been when developing relationships outside of academia bubbles and being with people with lived experience. • Having more, like, role-playing kind of activities might be great because for me, it's like if I'm in a situation where I have to speak up, my mind goes blank.
Digital demo module experience	• Enjoyed digital module (8 mentions) • Clear and organized web-based structure (6 mentions) • Particularly liked CBT-based examples of core beliefs (4 mentions) • Particularly liked online daily self-reflection logging (4 mentions)	• I really like the module. That's just like what happens at the hospital. • It was clear and I thought the structure was very helpful and consistent while going through the four examples of core beliefs • The good clinician one in particular led me to think about how there are so many ways the system rewards not thinking and not challenging biases, and I think it was nice that you provided that example.
Digital demo module time	• Feels that 15 minutes/week of web-based intervention content for six weeks is manageable (11 mentions)	• Yes, I think we can definitely make that time.
Demo module digital format	• Prefers online content with addition of in-person or group setting (7 mentions) • Concerns about exclusively online, self-directed formats (5 mentions) • Prefers self-directed online-only modules (4 mentions) • Open to conducting over a Zoom call (3 mentions)	• In-person is generally always the most effective. I think we tend to have short attention spans, and it becomes just an online module you have to do. • If you really want people to be an active participant and really engage with it, I don't know how good self-directed modules are … I'm just like clicking through it.
Digital demo module improvements	• Enhance resident participation and engagement in format (4 mentions) • Include web-based incentive to track growth (2 mentions)	• I think that it might be helpful to allow us space to bring up our own examples, but I know that it takes a lot of vulnerability for us to sit there and reflect. • If there's some sort of incentivization structure for people to check back in or record progress into, like a diary, I think that could be effective.
Potential report of EHR bias	• Yes, it would be helpful (11 mentions)	• Yeah, it would overall. I think it would be cool, because in the same way that they make us look at how often we are prescribing benzos, why can't we also be explicit, you know. in terms of antiracism? • Yeah, it would make me more wary. It would make me sit down and think.

XSL•FO
RenderX

Question	Themes	Quotes
Across questions	• Minority tax (5 mentions)	• Everyone has the responsibility to care for, like, a language diverse community. It [shouldn't] just fall on certain individuals just because of their background. • I remember there were some moments in medical school where I felt like there was too high a burden on faculty of color and also students of color.

Discussion

Overview

Using semistructured interviews with mental health professionals, our results indicate favorable acceptability of antiracist intervention content and digital delivery methods. The web-based demo module of the antiracism intervention received a high level of positive feedback, with participants finding it relevant, well-structured, and generally effective in teaching CBT principles. For example, participants enjoyed learning how to identify, react to, and consciously correct core beliefs that propagate racism in health care. Regarding acceptability, participants felt the time commitment would be feasible, especially the convenience of being able to access web-based modules for short periods of time over the course of several weeks. Online self-directed training was well-received, with a recommendation for the addition of a group, in-person, or zoom component to solidify and expand upon web-based self-directed learning. Participants also felt that this would improve engagement, especially with opportunities to share their own experiences. Such findings are in line with previous research suggesting that personalization and increased social connectedness facilitated by digital health intervention components can enhance user engagement [10].

In the context of existing literature, there is a need for targeted evidence-based antiracist strategies addressing the unique and specific needs of clinicians operating in any given health care specialty, as the needs of most mental health professionals will differ greatly from those of general health practitioners [11]. Unfortunately, most antiracism interventions to date have focused only on general health professionals, resulting in the existence of far less tailored interventions addressing a specific health care context or specialty. Furthermore, there are limited discussions of methods for enhancing engagement in antiracism training other than mandating antiracism work [11]. Findings from this study fill this critical gap in the literature by investigating needed aspects of antiracist intervention dedicated to specialized mental health care, with the added benefit of using novel digital-based design elements promoting enhanced acceptability and participant engagement.

Limitations

Limitations include the fact that this study was conducted at a single academic center, which limits its generalizability to other institutions. However, this is a targeted approach to be applied to the study population of mental health professionals. A similar approach can be applied to other health specialty areas, using interviews targeting clinicians of interest. Such methods may further be used to tailor digital antiracism training to other clinical specialties. Another limitation is that the current study focuses on the acceptability of the intervention rather than its efficacy. Lastly, another important limitation lies in the lack of community engagement in the intervention design process, an aspect known to enhance the health equity of digital health interventions [12,13]. Future iterations will therefore aim to involve the systematic incorporation of the voices of community members served by mental health professionals.

Conclusions

Taken together, these results provide important guidelines for the implementation of a targeted intervention for mental health clinicians. They suggest favorable acceptability regarding the use of CBT principles in antiracism education and delivery in a web-based format. Such synthesized findings and insights from mental health professionals may be used to tailor and guide practical aspects of the further development and piloting of a future large-scale web-based antiracism intervention.

Acknowledgments

Funding for this project was provided by the University of California, Los Angeles Initiative to Study Hate (TRB), the University of California, Los Angeles Department of Psychiatry Justice, Equity, Diversity, and Inclusion Research Support Funds, and the American Psychiatric Association Substance Abuse and Mental Health Services Administration Minority Fellowship program (#H79SM080388; TRB). The contents of this publication are solely the responsibility of the author and do not necessarily represent the official views of the Department of Health and Human Services, APA, or APA Foundation.

Data Availability

The data used for this study are available from the corresponding author upon reasonable request.

Conflicts of Interest

None declared.

XSL•FO
RenderX

References

1. Li H, Eack SM, Montrose DM, Miewald JM, Keshavan M. Longitudinal treatment outcome of African American and Caucasian patients with first episode psychosis. Asian J Psychiatr 2011;4(4):266-271 [FREE Full text] [doi: 10.1016/j.ajp.2011.08.004] [Medline: 23051160]
2. Bailey ZD, Krieger N, Agénor M, Graves J, Linos N, Bassett MT. Structural racism and health inequities in the USA: evidence and interventions. Lancet 2017;389(10077):1453-1463. [doi: 10.1016/S0140-6736(17)30569-X] [Medline: 28402827]
3. Jansen MO, Brown TR, Xu KY, Glowinski AL. Using digital technology to overcome racial disparities in child and adolescent psychiatry. J Am Acad Child Adolesc Psychiatry 2022;61(10):1211-1217 [FREE Full text] [doi: 10.1016/j.jaac.2022.03.013] [Medline: 35358663]
4. Brown TR, Xu KY, Glowinski AL. Cognitive behavioral therapy and the implementation of antiracism. JAMA Psychiatry 2021;78(8):819-820 [FREE Full text] [doi: 10.1001/jamapsychiatry.2021.0487] [Medline: 33950160]
5. Cénat JM, Broussard C, Jacob G, Kogan C, Corace K, Ukwu G, et al. Antiracist training programs for mental health professionals: a scoping review. Clin Psychol Rev 2024;108:102373 [FREE Full text] [doi: 10.1016/j.cpr.2023.102373] [Medline: 38232574]
6. Chu W, Wippold G, Becker KD. A systematic review of cultural competence trainings for mental health providers. Prof Psychol Res Pr 2022;53(4):362-371 [FREE Full text] [doi: 10.1037/pro0000469] [Medline: 37332624]
7. Friis-Healy EA, Nagy GA, Kollins SH. It is time to REACT: opportunities for digital mental health apps to reduce mental health disparities in racially and ethnically minoritized groups. JMIR Ment Health 2021;8(1):e25456 [FREE Full text] [doi: 10.2196/25456] [Medline: 33406050]
8. Lyon AR, Munson SA, Renn BN, Atkins DC, Pullmann MD, Friedman E, et al. Use of human-centered design to improve implementation of evidence-based psychotherapies in low-resource communities: protocol for studies applying a framework to assess usability? JMIR Res Protoc 2019;8(10):e14990 [FREE Full text] [doi: 10.2196/14990] [Medline: 31599736]
9. Braun V, Clarke V. Using thematic analysis in psychology. Qual Res Psychol 2006;3(2):77-101. [doi: 10.1191/1478088706qp063oa]
10. Borghouts J, Eikey E, Mark G, De Leon C, Schueller SM, Schneider M, et al. Barriers to and facilitators of user engagement with digital mental health interventions: systematic review. J Med Internet Res 2021;23(3):e24387 [FREE Full text] [doi: 10.2196/24387] [Medline: 33759801]
11. Hassen N, Lofters A, Michael S, Mall A, Pinto AD, Rackal J. Implementing anti-racism interventions in healthcare settings: a scoping review. Int J Environ Res Public Health 2021;18(6):2993 [FREE Full text] [doi: 10.3390/ijerph18062993] [Medline: 33803942]
12. Brewer LC, Fortuna KL, Jones C, Walker R, Hayes SN, Patten CA, et al. Back to the future: achieving health equity through health informatics and digital health. JMIR Mhealth Uhealth 2020;8(1):e14512 [FREE Full text] [doi: 10.2196/14512] [Medline: 31934874]
13. Nguyen-Truong CKY, Waters SF, Richardson M, Barrow N, Seia J, Eti DU, et al. An antiracism community-based participatory research with organizations serving immigrant and marginalized communities, including Asian Americans and Native Hawaiians/Pacific Islanders in the United States Pacific Northwest: qualitative description study with key informants. Asian Pac Isl Nurs J 2023;7:e43150 [FREE Full text] [doi: 10.2196/43150] [Medline: 36648292]

Abbreviations

CBT: cognitive behavioral therapy

Edited by D Gooch; submitted 07.09.23; peer-reviewed by K Xu, A AL-Asadi; comments to author 02.01.24; revised version received 04.02.24; accepted 06.02.24; published 03.04.24.

Please cite as:
Brown TR, Amir H, Hirsch D, Jansen MO
Designing a Novel Digitally Delivered Antiracism Intervention for Mental Health Clinicians: Exploratory Analysis of Acceptability
JMIR Hum Factors 2024;11:e52561
URL: https://humanfactors.jmir.org/2024/1/e52561
doi:10.2196/52561
PMID:38568730

©Tashalee Rushell Brown, Habiba Amir, Drew Hirsch, Madeline Owens Jansen. Originally published in JMIR Human Factors (https://humanfactors.jmir.org), 03.04.2024. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction

XSL•FO
RenderX

in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on https://humanfactors.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Validating the Effectiveness of the Patient-Centered Cancer Care Framework by Assessing the Impact of Work System Factors on Patient-Centered Care and Quality of Care: Interview Study With Newly Diagnosed Cancer Patients

Safa Elkefi[1], PhD; Onur Asan[2], PhD

[1]School of Nursing, Columbia University, New York, NY, United States

[2]School of Systems and Enterprises, Stevens Institute of Technology, Hoboken, NJ, United States

Corresponding Author:
Onur Asan, PhD
School of Systems and Enterprises
Stevens Institute of Technology
1 Castle Point Terrace
Hoboken, NJ, 07030
United States
Phone: 1 4145264330
Email: oasan@stevens.edu

Abstract

Background: Patients with cancer who have recently been diagnosed have distinct requirements compared to cancer survivors. It is crucial to take into account their unique needs to ensure that they make informed decisions and are receptive to the care provided.

Objective: This study suggested a framework titled Effectiveness of Patient-Centered Cancer Care that considers the needs of newly diagnosed patients with cancer and related work system factors. This study investigated how work system factors influence the perceptions of patient-centered care, quality of care, and associated outcomes among newly diagnosed patients with cancer. Patient-centered care is defined in terms of workload and communication considerations, whereas the quality of care is assessed through indicators such as trust in physicians, satisfaction with care, and perceptions of technology.

Methods: This study used qualitative data collected through interviews with newly diagnosed patients with cancer (N=20) right after their first visits with their physicians. Thematic analysis was conducted to validate the 5 hypotheses of the framework, mapping the interactions among quality of care, patient-centered care, and work system factors.

Results: We found that workload and patient-centered communication impact the quality of care and that the work system elements impact the patient-centeredness (workload and communication) and the quality of care (trust in physicians, satisfaction with care, and perception of technology use).

Conclusions: Qualitatively validating the proposed Effectiveness of Patient-Centered Cancer Care framework, this study demonstrated its efficacy in elucidating the interplay of various factors. The framework holds promise for informing interventions geared toward enhancing patients' experiences during their initial visits after diagnosis. There is a pressing need for heightened attention to the organizational design, patient processes, and collaborative efforts among diverse stakeholders and providers to optimize the overall patient experience.

(JMIR Hum Factors 2024;11:e53053) doi:10.2196/53053

KEYWORDS

cancer; communication; trust; satisfaction; technology; workload; work system factors

XSL•FO
RenderX

Introduction

Background

Improving the quality of care (QOC), coordination, and quality of life are essential goals of chronic care [1]. Patient-centered care (PCC) is one of the approaches used to assure the primacy of the individual's health and life goals in their care management [1]. In recent years, the concept of having the person be the driving force in their health care decisions has evolved and gained momentum, and it is now largely considered the gold standard for health care worldwide [1,2].

The initial physician visits after a cancer diagnosis are a critical period in which patients face a range of challenges that can significantly disrupt their lives. Symptoms of the disease and the overwhelming decisions related to treatment can pose a threat to their physical, cognitive, and emotional well-being [3]. Patients often struggle to comprehend medical information and express frustration with prolonged waiting periods for prognoses and follow-ups [3]. This can lead to psychosocial concerns, including high levels of distress, emotional strain, uncertainty surrounding mortality, and disruptions to social life [3]. The cognitive and emotional burdens can be overwhelming, potentially leading to nonadherence to treatment plans [4].

PCC approaches are considered crucial for the delivery of high-quality care to patients. However, there is considerable ambiguity concerning the exact meaning of the term and the optimal method for measuring the process and outcomes of PCC [5]. Despite the concept's popularity in the past 30 years, there has been a slight argument of perspective in the literature about the definition of PCC [5]. It has been an evolving concept, originally presented by Balint [6], who described patient-centered medicine as understanding the patient as a unique human being, whereas for Levenstein et al [7], it is an approach in which the "physician enters the patient's world to see the illness through his eyes." In 1998, Delbanco et al [8] developed a self-described utopian vision for a patient-centered health care system called People Power. The relationship is supported by "computer-based guidance and communication systems." Don Berwick, a former administrator for the Centers for Medicare and Medicaid Services, has popularized the slogan Delbanco and his group adopted, "Nothing about me without me," acknowledging that PCC is not always evidence based. In his 2009 *Health Affairs* article, he emphasized that PCC relates to one's set of decisions and choices of circumstances and relationships in health care.

This concept has received increased attention since the 2001 Institute of Medicine report, "Crossing the Quality Chasm" [9], where health care quality and system-of-care improvement efforts were linked to the 6 core values: safe, effective, efficient, patient centered, timely, and equitable. Since then, myriad clinical, policy, and research initiatives have been launched to promote the study, advancement, and implementation of PCC. Research later presented 8 primary dimensions of the PCC model (respect of values, physical comfort, coordination and integration of care, information and education, access to care, involvement of family and friends, and transition and continuity) [10]. In 2015, the World Health Organization released its framework on "people-centered health services" [11], emphasizing a focus on a system that adopts individuals', careers', families', and communities' perspectives into a trusted health care system.

PCC frameworks have proved to change the behavior of patients with cancer as they successfully engage the patient by incorporating his biopsychosocial support system into care delivery and ensuring sustainable development [12]. Involving patients with cancer meaningfully in the processes and responding to their emotions as part of PCC adoption have been linked to better health outcomes, more trust, and better engagement of the patient in their care [13]. Thus, to evaluate the effectiveness of PCC initiatives, the cognitive perception of patients with cancer needs to be studied in relation to their behavior within the care settings (eg, trust, satisfaction, anxiety, and engagement). On the other hand, achieving high-quality care is a complex pursuit in any setting, especially for cancer care. Improving the patient journey requires an integrated system of care and productive interactions among many system levels. By understanding the work system components, the design and integration of tasks, technology, and clinical processes can be reviewed to better support the needs of individuals while optimizing system performance. A supportive work environment and a highly engaged workforce correlate with improved quality of PCC and hospital performance [14]. Case managers, navigators, quality officers, and administrators may track patient outcomes at the population level. A study conducted in 2017 on postdiagnosis treatment communication with patients with cancer highlighted the importance of coordination among specialists, primary care, and other people involved in the care processes with patients to deliver necessary care as problems in coordination can lead to fragmentation in health outcomes and processes. However, existing initiatives and care-planning processes face barriers to adoption and implementation. To sum up, tools and initiatives designed to improve health care delivery through PCC need to be inspired by systems engineering principles as recommended by the Institute of Medicine and the National Academy of Engineering to identify, develop, and sustain best practices informed by the needs of survivors, caregivers, clinicians, organizations, and communities [13].

Due to the complex nature of the health care system, it remains hard to provide patients with care that meets their expectations without accounting for the work system in which they are receiving the care services [15]. However, to our knowledge, no framework focuses on PCC from a systems perspective. Human factors engineering interventions need to take into account issues across the whole system (system approach) with macro-ergonomic considerations, including organizational factors, to be more likely to significantly impact QOC. The Systems Engineering Initiative for Patient Safety (SEIPS) model of work system and patient safety, for example, emphasizes the principle of "balance" and focuses on system interactions that need to be considered to make significant progress in health care quality, linking the work system factors to health outcomes [16]. In addition, although many studies have focused on the workload of physicians and staff, no study has focused on the workload of patients with cancer. In this qualitative study, we explored the impact of work system factors on newly diagnosed

XSL•FO
RenderX

patients with cancer's perceptions of PCC and QOC and the impact of PCC on the QOC outcomes among newly diagnosed patients with cancer following a suggested conceptual framework.

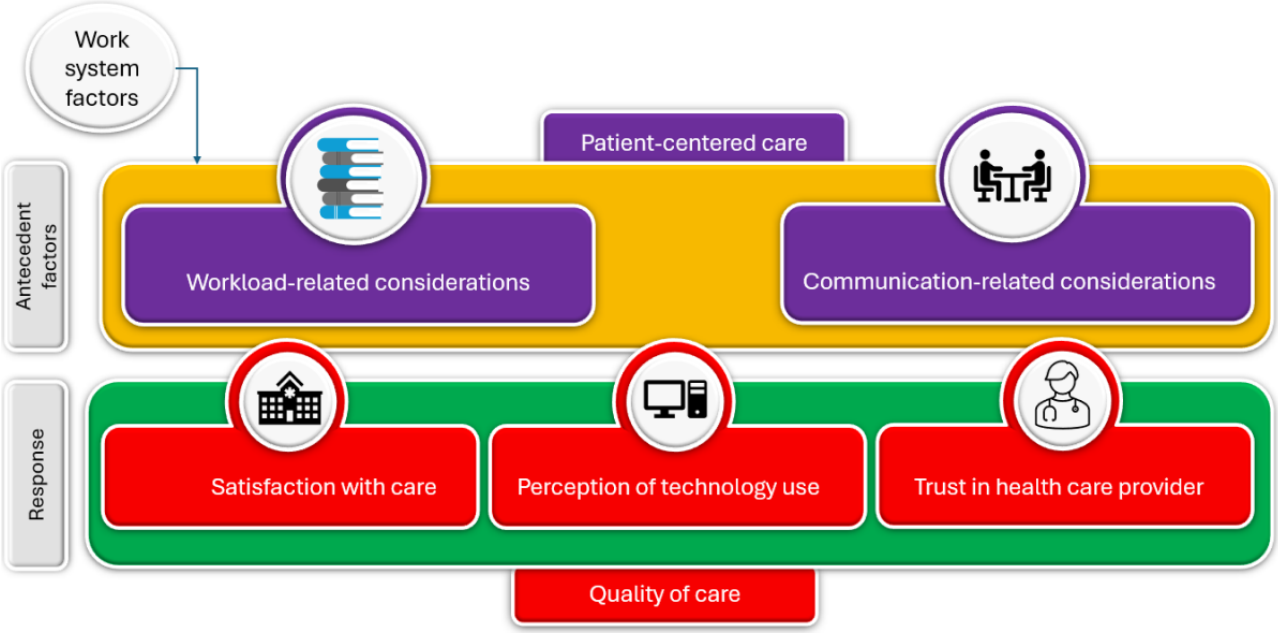
Theoretical Background

Overview

The framework built was inspired by different human factors models such as social cognitive theory [17], which conceptualizes the behavior of a person as a result of mental, personal, and social and environmental factors, therefore we considered behavior as a sum of a patient's perceptional cognitive input (patient-centeredness perception) and the response. Our patient-centered effectiveness components were inspired by the patient-centered communication in cancer care that defines communication through 6 functions [18] and the technology acceptance model. The technology acceptance model links technology perception to the attitude of the user toward the perceived usefulness and ease of use and the external variables [19,20]. The last model that inspired our framework is the SEIPS [21]. From a sociotechnical perspective, patients' experience, especially with chronic diseases, is a function of many coordination challenges [22]. Therefore, we need to go beyond the typical focus on a patient's single health care encounter and understand a patient's journey from a broader perspective through their interactions with other stakeholders in a system where not only patients and physicians are actors but the work system and the tools used are also important impactors of the perceptions and decision-making processes. Thus, we look at the systems' factors impacting patients' perception of patient-centeredness.

Our framework in Figure 1 emphasizes the relationship between patients' cognitive perceptions of patient-centeredness and QOC. We account for the impact of work system factors on these perceptions. We define patient-centeredness as a combination of workload support and communication and interrelationship support. Workload-related consideration characterizes the effective engagement of patients in their care experience. Communication and interrelationship improvement describes the communication effectiveness between patients and their providers. The dependent variables are related to the action tendency of patients: *satisfaction*, *perception of technology*, and *trust*. Exposure to the work system is considered a covariate in the model. To unpack this conceptual framework for evaluating patient-centeredness effectiveness, each independent variable has operational precedent in the human cognitive factors and behavioral economics literature.

Figure 1. Effectiveness of Patient-Centered Cancer Care framework.



Patient-Centeredness (Perceptional Cognitive Input)

Effective communication with patients with cancer can help meet information needs, improve physical and mental health, promote intimacy, and reduce burden [23]. In addition, patients diagnosed with cancer spend a lot of time and effort receiving treatment. Sometimes, patients have to deal with complex tasks related to medication taking and treatment in addition to rehabilitation activities that exceed their abilities, which engenders an overburden that has been proven to cause problems with adherence to treatment plans [24]. We define patient-centeredness in this conceptual framework as workload-related consideration and communication and interrelationship-related considerations.

Workload-Related Considerations (Ensuring Effective Engagement and Task Load Improvement)

Patient ergonomics is the application of human factors or related disciplines to study and improve patients' and nonprofessionals' performance of effortful work activities in pursuit of health goals [16,25]. A central emerging concept of societal views of health care considers that patients actively perform "work" to achieve health-related goals and objectives [26]. This way, human factors position patients at the center of the work system,

XSL•FO
RenderX

aiming to improve their experience with the workload assigned [25,27]. In highly sensitive situations such as cancer care, this paradigm can help us better understand the dynamics among the 3 actors of the visits (physician, patient, and technology) and how their interaction can influence critical outcomes such as QOC, trust of physicians, and acceptability and perception of technology use. We define the role of patient-centeredness as a booster to the effective engagement and performance of patients in their care through *task load improvement perception*. Thus, the effectiveness is measured through task load improvement. Cognitive task load or workload is used in human factors or organizational psychology. It operationally refers to the levels of difficulty that an individual encounters during the performance of a task and is a measure of human performance [28]. Subjective methods commonly used in research include rating perceived task difficulty, engagement, or effort made by research participants [29]. There are 3 types of workload measurement: physiological, performance based, and subjective [30]. The physiological workload measures concern the continuous size of the body's physical responses [30].

Communication-Related Considerations (Communication and Interrelationship Improvement)

Compared to other health care settings, communicating information during oncology visits, especially initial ones, is critically important but can be particularly challenging due to the substantial amount of information provided, complex treatment decision options, involvement of multiple different providers (surgical, medical, and radiation oncology), and highly emotional situation with high patient workload [31]. Patients might not recall information accurately and might face difficulties understanding the information given. When information is particularly upsetting, many patients are too stunned to register further information [32]. Patients report leaving initial visits feeling that their informational needs (particularly about treatment, side effects, and prognosis) are not always met [32], which can lead to uncertainty, anxiety, and depression [31]. In one study with newly diagnosed patient with cancer–oncologist dyads, agreement on the content of the topics discussed ranged from only 37.5% for treatment side effects to 60% for prognosis [33]. Incomplete or inaccurate information about the disease process and treatment options increases the likelihood of patients receiving a suboptimal QOC [34]. Misunderstanding resulting from lack of communication has impacted health care outcomes such as decision-making, trust, and effective treatment [35]. Many countries have opened their accreditation, certification, and quality improvement programs for the past decade to examine physicians in training and communication skills [36]. Interpersonal and communication skills are 1 of the 6 general competencies for physicians identified by the Accreditation Council for Graduate Medical Education and the American Board of Medical Specialties in the United States [37,38]. "While communication skills are specific tasks and behaviors performed by individuals, interpersonal skills are relationship-oriented and process-driven, as noted by Duffy and colleagues" [39].

Response (QOC Perception)

Emotional distress is an average expected reaction to a cancer diagnosis. The diagnosis causes psychiatric complications (eg, anxiety, stress, and depression) induced by the patient's perceptions of the stigma commonly attached to cancer [40]. However, it is widely recognized that patient-centered interactions have the potential to influence patients' behavior and well-being [41-45]. Thus, we model patient-centeredness as an influencer of the behavior, which is patients' perception of QOC (*satisfaction* with the care offered, *perception of health ITuse*, and *trust* in health care providers).

QOC Perception: Perception of Health IT Use

It has been long promoted that health ITs (HITs) will improve efficiency and QOC, support health care delivery, and reduce costs for the health care industry [46]. Much of the work has assessed how health care providers and organizations can use HITs to deliver health care services [47,48]. However, a growing awareness exists that consumers also want to participate in their health care [49]. For chronic disease settings such as oncology, patients must participate in the monitoring and managing of chronic diseases [50]. Several factors contribute to the widespread use of eHealth in chronic care; acceptance and capability of using ITs are vital components of understanding the disease and treatment options [51]. Advancements in digital communication and medical technologies have led to digitalizing health care [52,53]. The increasing adoption of various HITs has created new channels for physician-patient communication beyond the walls of physicians' offices. With the increased adoption and use rate of electronic health records in cancer care, oncologists can use the provided data in the critical decision-making process and support their workload [54]. In a study by Mazur et al [55], the enhancement of electronic health record systems' usability was associated with better oncologist cognitive workload and performance. However, little attention has been paid to technology support for newly diagnosed patients with cancer. Therefore, extending the existing knowledge base is essential to better understand how technology impacts newly diagnosed patients with cancer. Research on the mechanism of patient-centeredness shows that it is necessary to ensure patients' engagement with their health and their providers over the treatment time [56] as it impacts patients' lifestyles, quality of life, and behavior in the context of cancer care.

QOC Perception: Trust in Health Care Providers

Extensive literature supports the importance of trust in physicians for patients with cancer as it has been linked to improving QOC and other treatment outcomes such as adherence to treatment [57]. On the basis of a review by Hillen et al [57], trust is needed to ensure a good interaction between physicians and patients. Trust has also been shown to be impacted by communication among newly diagnosed patients with cancer [58]. Thus, we consider trust as one of the QOC factors affected by the communication and workload of newly diagnosed patients with cancer.

XSL•FO
RenderX

QOC Perception: Satisfaction With Care

Patients demand excellent care services from their providers. It is becoming a competitive edge in health care to control the quality outcomes and patients' satisfaction with the services, the providers, and the organizations in which they receive care [59]. Satisfaction is an outcome of utmost importance in cancer care [60]. It was shown to be related to physicians' ability to elicit the concerns of patients' with cancer, consider their psychosocial needs, and involve them in treatment decision-making, which are the techniques of "patient-centered" care and communication [60,61]. We consider satisfaction with care to be the third main component of the framework.

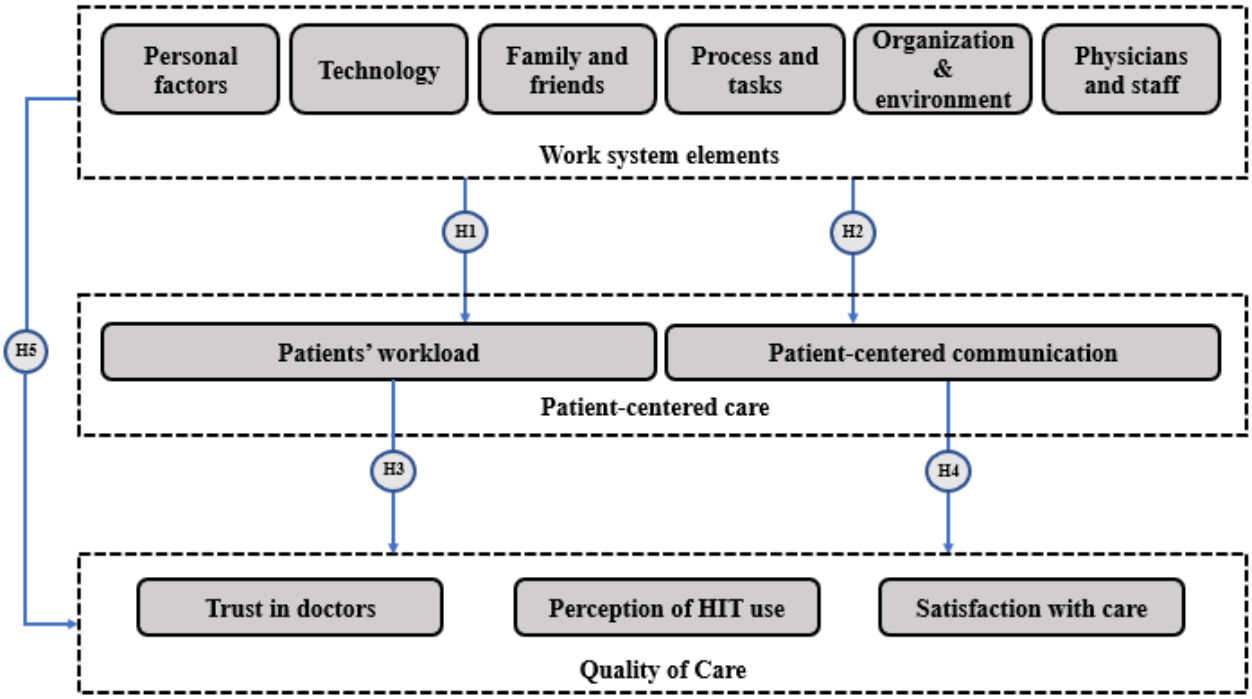
Hypotheses and Tests

Effective communication can prevent lapses in QOC and can mitigate harm when problems occur [62]. In cancer care, it is even more important to provide patients with the suitable communication needed [63]. Improving communication with patients with cancer in the first few visits requires a better understanding of patients' experiences of breakdowns in care and their needs in the early stage of their experiences [64]. In addition, patients frequently experience high load and feel overwhelmed due to their confusion about the treatment plans and their uncertainties about their options, which compromises their perception of QOC [34].

The complexity of cancer care, typified by the financial, emotional, and physical challenges, makes patient care challenging [65,66]. In addition, the complexity of the cancer care work system is reflected in the multiple clinicians that are involved in the processes, the long therapies, and the uncertainty of the outcomes [66]. Thus, we considered the following hypotheses: (1) work system elements—work system factors impact newly diagnosed patients with cancer's perception of their workload (hypothesis 1); (2) work system elements (communication)—work system factors impact newly diagnosed patients with cancer's perception of their communication with physicians (hypothesis 2); (3) workload (QOC)—workload impacts newly diagnosed patients with cancer's perception of QOC (trust in physicians, satisfaction with care, and perception of HIT use; hypothesis 3); (4) communication (QOC)—physician-patient communication impacts newly diagnosed patients with cancer's perception of QOC (trust in physicians, satisfaction with care, and perception of HIT use; hypothesis 4); and (5) work system elements (QOC)—work system factors impact newly diagnosed patients with cancer's perception of QOC (trust in physicians, satisfaction with care, and perception of HIT use; hypothesis 5);

We consider the following as work system factors: physicians and staff, organization and environment, family and friends, and processes and tasks. All hypotheses are summarized in Figure 2.

Figure 2. Hypothesis framework guiding the qualitative analysis. H: hypothesis; HIT: health IT.



Methods

Overview

This study used qualitative data from semistructured interviews to explore the impact of work system factors on newly diagnosed patients with cancer's perceptions of PCC and QOC, as well as the impact of PCC on QOC and technology preferences. We used semistructured interviews to facilitate candid disclosure of personal experiences. The interview questions used in this study were guided by the SEIPS 2.0 mode [26] and validated by existing literature. The full interview guide has been published elsewhere [67]. The SEIPS 2.0 model provides a framework that helps comprehend the work system (people, tools and technologies, tasks, working environment, and organization); process (clinical process, patient outcome, and

XSL•FO
RenderX

organizational outcome) in the health care domain [26]. It also helps assess and understand the complex interaction among work system elements [26]. The interview guide was developed iteratively with the research team. Subsequent revisions of the interview guide were informed by emerging themes and sensitizing concepts generated through data collection and analysis. We also revised the interview guide questions based on expert feedback and the results of quantitative research conducted previously in the same center as part of the same project. Our sample included patients who (1) had been recently diagnosed with cancer, (2) were in their first few visits to the cancer center where the study took place, and (3) were adults aged ≥18 years.

We conducted a total of 20 in-depth semistructured interviews. Sampling continued until theoretical saturation was achieved, defined as the point at which further interviews did not advance the conceptual depth of the developed categories or reveal new dimensions of the relationships among categories [68,69]. The interviews lasted approximately between 30 and 60 minutes and were facilitated by a trained expert in clinical research management. The length of each interview was determined by the patient's level of comfort in disclosing their perceptions and sharing their experiences. We completed 20 interviews, resulting in 989 minutes of recording that were used for data analysis. All participants provided informed consent for the interviews to be audiotaped and professionally transcribed.

Analysis of interview transcripts was iterative and used a deductive and inductive approach. The deductive approach used focused coding, applying predetermined codes or themes resulting from the preset hypotheses made regarding the different interactions among the perceptions of work system factors, QOC, and patient-centeredness. Throughout the study, we incorporated memo writing to reflect on individual cases, interview settings, participants' responses, and emerging concepts and assess preconceived notions that were discussed weekly with the research team. The coding was done and visualized using a Microsoft Excel (Microsoft Corp) spreadsheet. We also prepared the COREQ (Consolidated Criteria for Reporting Qualitative Research) checklist and provided it in a supplementary file (Multimedia Appendix 1).

To mitigate the risk of bias caused by qualitative research, our study initially used a triangulation design with findings reported in different studies [3]. For more transparency and accuracy, different participants with different backgrounds reviewed and confirmed the transcripts and interpretations. In addition, a clear documentation of the analytical process was conducted.

Ethics Approval

This study received ethics approval (institutional review board ID 00011536) from the Stevens Institute of Technology and from Hackensack Meridian Health, John Theurer Cancer Center, New Jersey, where it took place.

Results

Overview

The distribution of demographics is shown in Table 1. Most participants were female (12/17, 71%), White (9/17, 53%), and aged >40 years (15/17, 88%).

Table 1. Sociodemographic characteristics of the study sample (N=17).

	Participants, n (%)
Gender	
Male	5 (29)
Female	12 (71)
Race	
Black	2 (12)
Hispanic	4 (24)
White	9 (53)
Other	2 (12)
Age (years)	
18-39	2 (12)
40-59	10 (59)
≥60	5 (29)
Education	
High school or lower	7 (41)
Bachelor's degree level	5 (29)
Graduate school	5 (29)

XSL•FO
RenderX

Impact of PCC on QOC

This section presents the findings related to the testing of hypotheses 3 and 4.

Workload of Newly Diagnosed Patients With Cancer

Newly diagnosed patients with cancer expressed their perception of the workload and reported experiencing a high mental workload due to frustration and emotions when they were first diagnosed:

> *At first you're all nervous and upset with your condition. So you're like, I don't know how to do this. You're freaking out.* [Patient 13]

The mental workload increases as patients are required to remember many details (eg, appointments, information, and medication) and are required to understand the results and options they are given with the little information shared with them:

> *There's a lot to remember. There's a lot you got to remember which medicine to take and when.* [Patient 12]
>
> *I think sometimes the interpretation of the test results is challenging and a bit anxiety producing.* [Patient 06]

In addition, patients experience high temporal and physical load as their tasks require a lot of effort and are time and energy consuming:

> *It's unfortunate that I have to get drawn so many time...You just get tired of sitting, you get tired of being in there.* [Patient 02]
>
> *Very demanding. I go get the blood work done first, and then when my blood work is done, I go up and wait to see the doctor. Then I go to the infusion room when I'm done with him.* [Patient 13]

Impact of Workload on Trust in Physicians

High workload negatively impacts newly diagnosed patients with cancer's trust in their physicians. For instance, patients who had less workload, felt cared about, were given enough time, or were less rushed were more likely to trust their physicians:

> *My doctor really care about me. I gage that because I'm not being rushed to leave right away. I see the attention that they pay attention to all my questions that I have, and also, I feel comfortable.* [Patient 03]
>
> *I think they do help with my emotions because they are trying to address every question that I have. I am not getting very emotional; I'm not getting extremely upset. So what I'm trying to say is that they might not even see any emotional reactions from me because I do trust that the doctor said he will fix me and I know that he will. So, I'm not emotional. I'm going very strong because of their support and my knowledge that they are there, and they will fix the situation.* [Patient 05]

Impact of Workload on Satisfaction With Care

The findings also showed that the workload perception among newly diagnosed patients with cancer impacts their satisfaction with the care services received:

> *As far as the demands, the time and effort it took to ensure the entirety of the visit was completed well in its entirety. Look, it was very effortless on my part.* [Patient 06]
>
> *They just gave you time to digest and let you just sit there and think about what the doctor was saying, and that really does make a difference. So basing out the information across the appointment instead of just kind of ramming it all down your throat at the beginning and asking any questions, it always moves quickly. But I do think that the taking of time in between providing information and providing time to digest was very helpful.* [Patient 12]

However, patients who felt that the care services were demanding of time or effort were less satisfied with their visits:

> *The treatment itself, when I'm there for chemo, does take a long time. That's like 4 hours at least. And then I'm on a pump for the next two days, so that does take a decent amount of my time. And then when I'm on the pump, I'm basically laying down because I really can't do much because I'm still getting treated. So, the treatments are demanding.* [Patient 13]

Impact of Workload on Perception of Technology Use

Finally, the high workload impacts patients' approval of physicians using technology during the visits. Patients who felt that using technology made their tasks less effort, memory, and time consuming were more likely to accept it. They felt that having all the information available in one place saved them from looking for information everywhere and trying to understand it and share it again with the physicians:

> *Technology can help keep me cooperating with the doctors and the treatment without much effort...Usually the nurse has the computer and she's checking in that all the details are correct and verifying information with me before the doctor even walks into the room. And she can answer questions about when I start the next cycle, et cetera. Because she has all the information with her. That saves me time. I don't need to remember everything.* [Patient 02]
>
> *It will impact only positively because I see an order, everything in one place in my app. I see my history, I see all my tests, I see all my appointments, all notes from the doctor, I see all the scans and everything is in one place. For me it's positive because again, I'm technology savvy and for me, that's great to be utilized, to have everything in one place.* [Patient 05]

To conclude, these qualitative findings showed that having less workload (eg, physical, temporal, and effort based) helped newly diagnosed patients with cancer trust their physicians more, be

XSL•FO
RenderX

more satisfied with the care visits, and accept technology use during the visits (hypothesis 3 was supported).

Communication Needs and Impact of Communication on Trust in Physicians

The quality of the communication in cancer visits is critical. Physicians share a lot of information with patients in the first visit, and it is essential for them that physicians explain things clearly in an easy, comforting, and understandable way and that their questions are answered without being rushed. Good communication of the information needed made patients trust their physicians:

> *She made it very easy for me to understand very difficult, very complex procedures and explain to me clearly why, in my particular situation, she would recommend the clinical trial treatment that I'm on. And the appointment didn't last more than maybe 20 minutes, but in that time, I felt like I understood what I was getting myself into. She provided things for me to read when I got home and explained clearly that other options were not as good as this one...If I was concerned about something, she would try to reinforce that. Everything that she explained made sense because she was trying to keep me at ease and not worried about more than I should have.* [Patient 02]

> *Our communication me and the doctor is genuine, accessible, nurturing, informative.* [Patient 17]

Different patients also have different levels of understanding. Patients trust their physicians if they respect their pace and health literacy levels:

> *Yeah, basically pausing and making sure I understand things as she's saying them, making sure I'm caught up in the discussion, because you could go very fast over a lot of things and there's a lot of information to digest at an appointment like that. So she would take her time to pause and say, do you understand what I'm talking about? So that sort of thing helped.* [Patient 17]

Another important communication factor that impacted patients' trust in their physicians was feeling like they were treated equally without any bias independent of the type of insurance they had or their age or race. Patients liked to be treated as human beings, not as numbers:

> *It felt good. I didn't feel like I was going to be just thrown in there and just done whatever, and then no explanation of anything, so it felt really good. And the mere fact that I have Medicaid and I didn't feel like a patient, that felt phenomenal for me. It was very powerful. Well, I mean, it seems that we get lost in the shuffle. We're just like, nobody basically. We seem to be like, I don't know, nobody, we're not treated like we just don't matter. I didn't feel that way at all. I felt like I have a chance just like everybody else who has like [insurance name] or whatever. So I just felt like they cared about me and the process was they were going to do whatever they had to do, no matter what insurance I had. So that meant a lot to me.* [Patient 07]

Furthermore, newly diagnosed patients with cancer like to be given full attention by their physicians, communicating both verbally and nonverbally, exchanging eye contact, and paying enough attention to every question they ask. This impacts their trust in physicians:

> *Well, they always examine me, obviously. They talk to me. They come up to me, and they look at me not everywhere in the room, just looking at me eye to eye, and explain to me exactly what happened this week...They look at me straight in the eyes, and for the time that we're there, her attention is focused on me. And when I ask the question, they usually don't mind repeating themselves, because sometimes when you're in treatment, you don't hear well. And I'm taking notes when I'm there. And I sometimes repeat questions that she may have already answered. And she is very happy to follow up and expand a little more so that I can understand in more detail what she's trying to tell me about follow up questions, answering follow up questions.* [Patient 02]

Finally, patients want to be treated in a personalized way as a special person and to build a strong relationship based on empathy with their physicians by talking about their personal life and not only about treatment and visits:

> *My doctor is extremely approachable despite his busy schedule. He shared his cell phone number with me, but of course I'm not going to communicate to him. But those are things that you understand that you're not just a patient, you're a special individual for him and for his staff, and everything is personalized. I think that's my belief, honestly.* [Patient 05]

> *I like that the doctor talks to me in general not only about my treatment. Sometimes we talk about family and things like that, really getting to know the doctor. He was very communicative; he keeps being positive and that helps.* [Patient 01]

> *I think in particular, she has quite a good amount of empathy, which I think a lot of doctors don't. So she treats you like a human being and trying to think what else on the medical side. It's basically kind of explaining your condition well and giving you an idea of what's to come.* [Patient 10]

In fact, patients who do not have a strong bond with their physicians and only talk about treatments without discussing their options may have low trust in the physicians' opinion and feel that they are treated in an unfair way:

> *Well, yeah, I feel that, but I think this is more and more interested in selling products that are profitable for the hospital than necessarily what care I need. She just seems obsessed with selling an expensive procedure that I'm not ready for. I'd like to see more programs tailored to my situation and some options. There's a lot of options and treatments available*

XSL•FO
RenderX

today. And to say it's this or nothing, I don't think it's appropriate. [Patient 11]

Impact of Communication on Satisfaction With Care

The communication between newly diagnosed patients with cancer and their physicians also impacts their satisfaction with the care received during the visits. In fact, patients prefer to be told the truth about what is happening to them and for it to be stated that the physicians are doing their best:

> *Well, anything the doctor noticed, any concerns she has, she always meant that she just wants to be sure and just saying that she wanted to stay on top of things and that was pretty good to me. I would say I had really good care both in first visit and follow-up ones.* [Patient 01]

They also want physicians to explain the goal of the visit clearly and to be walked through every step of the procedures. On the basis of patients' needs, physicians need to ask them whether they are satisfied or not and allow families to be part of the visit for more support:

> *When I have to come visit, they know exactly what happened before, and they can specifically tell me where we are today and what we need to do today. They also keep some time for me to ask questions, and if I don't understand something that she explained, she'll explain it again. My daughter participates in those visits as well, and she always has questions too. As long as we have questions. If we are all clear and everything's just a routine visit, then we need less time, I guess. But if we need more time to answer questions, they're willing to answer until we're satisfied...I like it when they explain things to me, even if they're technical, because I can look it up and kind of make sure that I'm getting the right information from my doctors and it's consistent with what the best health organizations treating cancer are doing.* [Patient 02]

Although some prefer to be told everything, others do not understand much when it is too technical. Physicians need to explain things to them in an understandable way:

> *The communication was excellent. The answers that I was seeking were given and the questions that I had were answered. The care was given, felt comfortable. The doctor communicated with me in a way I could understand.* [Patient 06]

However, paying attention to special cases is important to gain patients' satisfaction. Some patients have comorbidities and need to be taken care of in a more careful way. For example, one of the patients whom we interviewed was blind. It is more challenging to communicate with such patients:

> *I am blind, so they always print documents out for me though but they tell me everything verbally. To make sure I understand and then they tell me it's on the printout.* [Patient 16]

Some other patients are skeptical about health care systems. It is important to know how to handle them and how to communicate information to them without losing their trust not only in physicians but also in the system itself:

> *Cancer care is a profit center for these medical centers. The doctor is trying to push a very expensive procedure that's very invasive that I don't have the support network to do. And it seems to be like she wants an all or nothing for that procedure. So, I think I really need a second opinion on this stuff. Well, like I said, I think that the cancer centers seem to be out to maximize profits because I see them advertised all over. I don't know, that's just what I seem to find out.* [Patient 11]

Finally, despite the focus on visits being very important, follow-up needs to cover home care for the first visits as patients need more support at the beginning of their experience and building a bridge of communication with them beyond the care visits would help them feel more cared about:

> *So, the communication during the treatment and while I am in the hospital is really good and I feel that I can ask any question and I always get the answers. As for communication when I am at home, I think I am still learning the system. After my first treatment, I had some adverse effects. I did document and I did write up my observations. Science in me did that. But I didn't know how to communicate that to the doctor until my next follow up with the doctor. And then we did discuss those adverse effects and they did adjust my dosing regimen.* [Patient 05]

Impact of Communication on Perception of Technology Use

Newly diagnosed patients with cancer need their physicians to communicate with them without any distractions:

> *No, no them using a computer is of no distraction at all. The doctors still attend to me.... They always check my blood work and put it on the computer and if there is anything they communicate after the visit. They always call.* [Patient 01]

If patients are made the center of the visits and the computer is used for documentation purposes, patients feel satisfied with its use by physicians:

> *I don't think technology is disturbing my communication because they are still there. It's not that we are communicating through computers only. They are still there in the room they are personally discussing with you. But then they document everything in the computer. And I think at this time and era, you do expect that everything will be documented on the computer? That's my expectation.* [Patient 05]

> *She occasionally doesn't always use the computer, but occasionally does to look at test results. But I never really found it distracting, and I don't feel like she was paying attention to the computer more than me. It was just there as a tool as part of the*

XSL•FO
RenderX

> *appointment. Never did I feel like it was computer first, patient second sort of thing.* [Patient 12]

However, if the use of technology made patients feel that they could not communicate well with their physician, they were not happy with it being used during the visits:

> *Not every time, but yeah it was distracting, sometimes. Actually, now that I think about it, I think that was where I could see a few instances where that was where their focus was. I think I would ask a question and there would be like a two-minute pause because he was in the middle of typing stuff on the computer and then he would answer after. So, the doctors was distracted with whatever he was doing on the computer.* [Patient 08]

> *I would prefer them not to be on their computer and rather making eye contact and communicating directly with the patient rather than typing. So I didn't feel as connected with the doctor.* [Patient 17]

To conclude, these findings showed that, if the communication between physicians and patients is built in such a way that patients are the center of the care and using technology does not distract physicians from building a bond with their patients, technology use during the visits is accepted and not judged as distracting. Thus, hypothesis 4 is supported.

Impact of Work System Factors on Communication

The work system factors in health care impact patients' communication with physicians. In fact, patients were more satisfied with the communication with health care staff when they felt that the organization was empowering nurses to intervene and raise issues related to their health. Thus, the *organization and environment* impact the communication perception among newly diagnosed patients with cancer:

> *There was a nurse, also a night nurse, who noticed that there was fluid in my lung and she put a note in for the doctor to see if they could remove it because she thought it was too big. The day after I got in from the emergency room and the nurse was the one who raised the flag to the doctors and the next day they removed the fluid. So I think they're empowered, but at the same time looking she didn't have to go back and look at the X rays in my lung because she was surprised that I was breathing so badly. So, she was just curious and checked in and brought it to everyone's attention. I also like that they called me directly to check on me. They make sure that we are cared about and that we know everything about our situation.* [Patient 02]

In addition, frequent follow-ups can help patients share their concerns and issues with their physicians and help them communicate well in a continued way, which shows the impact of the *processes and tasks* on communication for newly diagnosed patients with cancer:

> *So, I think follow-ups are very important, especially after the first treatment. When my first treatment was very miserable, I felt very miserable afterwards, I had very significant adverse effects to the drugs. And then in a week I had follow up with the doctor and we had really good communications for the second treatment.* [Patient 05]

The way in which physicians, nurses, and staff interact with patients impacts their perception of communication. For instance, patients were more satisfied with physicians and other staff if they felt cared about and if their questions were answered. In addition, allowing family members to attend visits may help patients feel more reassured:

> *The people in the lab are amazing. They understand that we get pinched a lot, and they try to work with you, and they help each other, too, because they have to get the results stat, how they like to say. And they look thoroughly at the request from the doctor. And if I have questions about what they're doing, they'll answer them intelligently...When I first saw my doctor, she knew the record just as well, but she asked me to tell her my experience so she would know firsthand from me how I was feeling now and what had happened in the past. So, I felt very well taken care of and the communication between my doctor and me was excellent really.* [Patient 02]

> *My daughter participates in those visits as well, and she always has questions too.* [Patient 02]

Thus, we conclude that the work system elements impact the perception of newly diagnosed patients with cancer regarding communication, which supports hypothesis 2.

Impact of Work System Factors on the Workload

The work system elements impact patients' workload. In fact, the process of detailed documentation in the records and providers accessing that information easily also reduces patients' mental workload and frustration. Patients also like being guided through every step at the clinic as it helps them feel better:

> *I do feel supported, even though we meet for short periods of time...I felt that in every visit, in the few minutes that she was maybe 15 minutes that she's in the room, she knows everything about what happened the previous weeks. I don't know how she does it, but if I forget to tell her about something that I was feeling the week before, she would ask me about it. So, I understand. However, they do it to go into the room and remember exactly this patient in particular, it makes me feel very reassured that they've done their homework when they walk into the room to talk to me.... I think they have a pretty good system. Once you register, they have someone already greeting you, walking you over to do any lab tests that you need. They kind of wait around and guide you to the elevator so that you can go up to the waiting room waiting area.* [Patient 02]

> *You come and you get greeted by a person that sits on the first floor. Then you go to lab. In the lab everyone is very attentive. Sometimes you have to wait a couple of minutes, but usually it's not very long wait time and they are very attentive to ensure that they are doing very good.* [Patient 05]

XSL•FO
RenderX

However, the long waiting time in the process makes patients anxious and nervous, which adds more workload to the physically demanding processes and procedures that they are experiencing:

> *I just wait in between seeing the nurse and the doctor a little bit too long, I thought. So, the wait was a bit lengthy, a bit long. That was nervous for me. And not only that, but we had to leave at a certain time because we had to go pick up my nephew after school. So that was my appointment was at I think it was at 110, and I didn't see the doctor to, like, almost 230 or something like that.* [Patient 07]

> *I sat there, and I waited, and I waited, and I waited for my first biopsy results and to get them at 02:00 p.m. I was calling and calling and calling, trying to see if anything came in...It's a lot and it takes a toll on you.* [Patient 13]

To reduce this load, physicians and staff members need to explain matters clearly to their patients to comfort them, reassure them, and make them feel cared about through personalized services:

> *I would say the doctor knew how much information I needed to avoid being overwhelmed. Just telling me the options of what we need to do and I think that pretty much helped. Not feeling overwhelmed, it's like, let's do this and get past it. So that was pretty much my feelings.* [Patient 06]

> *I was very nervous about what the nurse had told me that was going to happen once. I didn't want to need to have a tube in my lungs. But luckily, before we got to the procedure, they had already taken care of that and she put it in capital letters so that the radiologist didn't miss it, that they didn't need to put it too, because the treatment would resolve that over time. So, I think reassurance is what they tried to do and being attentive to the details, which in medicine, I think is very important because each case is a different case. And I felt very comforted that I'm not just a number, I'm a patient that they're trying to get out of the hospital.* [Patient 02]

> *What I do when I am overwhelmed is I call the nurses all the time, and they're so helpful. I was calling them multiple times a week, and whether it was a new side effect, or I just had follow-up questions. So, I definitely have been utilizing them, and they've been so incredibly helpful.* [Patient 08]

In addition to the health care actors, the organization impacts patients' workload. Patients need a relaxing, calm, clean, and organized environment. Using comforting colors and decorative signs that motivate patients can also give them more hope and reduce their load:

> *So, when we look at the physical layout of it and all those processes, it's very nice, very organized place, very relaxing when you have to wait, so it's no problem.... Everything was very comfortable. No noise at all. Very calm and especially very accommodating.* [Patient 01]

> *The environment is really good. Everything flows nicely. Everything is nice and clean. Everything the colors and the walls and everything is very calming. As far as decorations and stuff like that, there's like a passageway that sometimes I'll pass through that I see like puzzles of past people that I have done from cancer. They do like these jigsaw puzzles, and they've hung them all across this hallway that I pass, which I find very endearing when I pass, and I see that. So pretty much in the sitting areas and everything where everyone sits and waits nice, valid and everything like that.* [Patient 03]

Construction work, long walking distances between the rooms, parking, and many other issues may cause patients to experience more physical and temporal load:

> *Let's be honest here. Like, it's completely overbooked, and I know they're under construction or what have you, which is stressful to have that many people crowding in the hallway.... Like if I was upstairs and I couldn't make it down in time, they would call and say, hey, she's running late. And it was accommodated. But it's definitely overwhelming to navigate.* [Patient 17]

> *You can build a bigger parking lot if you have room. Yesterday, yeah, I was riding because the appointment was at 01:00 at that time was all packed. They had to drive around all the way up to the roof and start coming down...there are definitely not enough spots there.* [Patient 09]

Furthermore, family and friends can help support patients in their tasks, which reduces their workload by facilitating processes:

> *My daughter is there for me. I moved in with her so that she could drive me to my appointments, and I can have support when I'm not feeling so good.... It makes things way easier.* [Patient 02]

> *Wife, family, friends, taking me to the treatment, stopping by and visiting me, phone calls. Just a lot of support.* [Patient 08]

> *My husband is always next to me. I am 90% self-cared, but sometimes I need his help to move around. For example, during the chemo to go to the bathroom.* [Patient 05]

On the basis of these findings, we showed that work system elements impact patients' workload. Thus, hypothesis 1 is supported.

Impact of Work System Factors on Satisfaction With Care

The work system elements impact patients' perceptions of QOC. They impact patients' trust in their physicians, their perception of HIT use, and the satisfaction with the care received during the visits. In fact, patients appreciated nurses checking on them

XSL•FO
RenderX

frequently and being nice to them, which made them feel more satisfied with the care received:

> *I think specifically the nurses in the infusion center, they were so kind and so nice, and they definitely were always asking how I was feeling during the infusions. They come check on me every ten minutes, pretty much. So, they were very accommodating and made me feel very comfortable.* [Patient 08]

> *Before seeing the doctor, you see the nurse nurses always welcoming even the staff that you go for copay or just approach to announce that you arrived. They are very attentive. You can see that they are feeling your pain. And that's very comforting, let's say...Usually my chemo is very long. So, the people who delivers lunch are so attentive to every single person. They are taking time for every single patient to repeat whole menu and convince you that this is very delicious to take. It's really warm and nice atmosphere. So I think they are taking extra steps to make you feel as comfortable as possible given the heaviness of the disease.* [Patient 05]

Receiving less attention from physicians would make patients unsatisfied with the visits:

> *My doctor is very busy. He's the head of the cancer center and he has tons of patients. But I would have liked to see him maybe sometimes not at the end of the treatment, but also in the middle of the treatment.* [Patient 05]

In addition, having a good organizational environment increases patients' satisfaction with care. Patients want to be in a well-designed environment where they can access better care services:

> *So, I think all is very accessible, very well designed, that you stop by first in the lab, then you can immediately pick up your pharmacy needs and then go to the second floor to visit the doctor and then move to the chemo center. So, I think everything was designed well. I love the sunny side and shady side of the infusion room. They have all these blankets, very nice people always asking what you want, more water, more anything to make you feel better. I think, as you say, from the organizational and structural perspective, is designed very well.* [Patient 05]

Impact of Work System Factors on Trust in Physicians and Staff

When physicians provide them with personalized services that are not based on generic information and that speak to their needs and situations, patients tend to trust them more:

> *It's just from reading the reports that the doctor gives me after the visit summary, I can tell that it's not generic. It's definitely speaking to my condition. I can see that what they're writing, and my evaluations are definitely about me, and I can see the reports, and it's definitely very much personalized.* [Patient 12]

The mix of personal and professional interaction makes them trust their physicians and the nurses delivering the services. Patients need people who listen to them, a friendly environment, and practices in which the main goal is to deliver the safest care to them:

> *What I did notice you have very good clinical practices that when the nurse has to introduce chemotherapy, they have second pair of eyes verification, which speaks of high quality and regulatory compliance of your organization, and that is incredible to see.* [Patient 05]

> *Every time I have a discrepancy, they always double-check. Either the nurse in the infusion room double-checks with the research nurse, the research nurse double-checks with the doctor, and everybody double-checks to make sure that we're doing it the right way, and that gives me comfort as well.* [Patient 19]

> *I think it's a good team. They listened to their own people, and they acted on it. That makes me trust the whole thing. So, I'm like I said, very professional and very personable. I really like that. It sums it up so beautifully. Like the two pieces in health care. Professional yet personable. I really do like that.... The hospital has great practice, and we have a number we can call. Twenty-four, seven. And they told us exactly how to behave and where to go when we got to the hospital. So that kept me more of these. But I was very scared. It was a Sunday night, so the doctor didn't physically come, but the emergency doctor had spoken to her. He knew my case.* [Patient 02]

> *They had a social worker contact me, which was nice to see if I had any opportunities or anything. I thought that was a good guess here. Any helping hand is an essential hand.* [Patient 11]

In addition, more trust is built among patients if the team's communication is healthy and professional:

> *The communication between the staff, I see that it is good and very professional. And the place is amazing. They walk with me and make sure I have all what I need to start the next step. The infusion, I prefer the one where I get sun and they already know that.* [Patient 02]

Impact of Work System Factors on Technology Perception

If physicians let themselves be distracted during the visits or do not pay enough attention to patients, patients may consider technology as a source of distraction and disruption to the visit:

> *The doctors was distracted with whatever he was doing on the computer.* [Patient 08]

Finally, the organized process of the visit, the good collaboration between nurses and physicians, and note-taking to make patients the center of the visits made patients trust their physicians more,

XSL•FO
RenderX

be more satisfied with the visits, and accept the potential that technology may have in the success of the care processes:

> *Doctor showed up even with nurse practitioner. And normally the doctor talks to you, she examines you, explains things. And then normally the nurse practitioner fills up whatever they need to do in a database. And normally it's not distracting at all. It's all adequately it happens in the background, and you concentrate on the conversation with a doctor, and someone else is filling out all the paperwork. Something needs to be like sending the prescriptions to my pharmacy and setting up another test. Everything was done at the same time, but I don't feel it was destructive at all. It was good.* [Patient 09]

Thus, to sum up, the health system elements impact patients' perception of technology use, trust in physicians, and satisfaction with care, which supports hypothesis 5.

Discussion

Principal Findings

In this study, we qualitatively explored the impact of work system elements on QOC and PCC and how PCC also impacts QOC among newly diagnosed patients with cancer in the first follow-up visits after the diagnosis. We found that newly diagnosed patients with cancer experience a high workload (mental, physical, temporal, effort based, performance based, and emotional) resulting from the frustrating diagnosis and the load of information that they receive in the first visits. This load impacted patients' trust in their physicians, satisfaction with care, and perception of technology use during the visits.

A diagnosis of cancer is a threat to one's sense of security, whereas feelings and emotions accompanying the disease uproot everyday existence [70]. Patients find themselves unpredictably facing a high emotional load and under the obligation to cope with the stress and anxiety caused by their diagnosis [70,71], which explains the high emotional and mental workload faced by our participants.

In addition to that, patients with cancer have to deal not only with the physical ailments resulting from the illness and its treatment but also with the thoughts of permanent health impairment, disability, fatigue, and pain that may result from their diagnosis [72], which correlates with our finding of high physical and effort-based workload perception among the participants. This may explain the dissatisfaction of patients with the quality of the care received. Emotional stress and mental problems can cause difficulties in everyday life, such as not being able to work, financial problems, and a lack of social support. This has been shown to impact quality of life perception among patients with cancer in other studies [73]. The literature also shows that patients with cancer can experience a variety of needs as each person reacts individually to the hardships of illness depending on their personality traits and understanding of their new situation [70]. With the substantial incoming flow of information, patients may find themselves unable to trust physicians and may consider technology as a distraction to their visits at that stage.

We also found that newly diagnosed patients with cancer can be very needy when it comes to communication with their physicians and that their communication with physicians impacts their perception of QOC. Communication is the cornerstone of the relationship with the patient in all medical settings, specifically chronic care, with the main aims of creating a good interpersonal relationship, exchanging information, and making treatment-related decisions [74]. Certain attitudes, behaviors, and skills (eg, ability to impart confidence, empathy, "human touch," relating on a personal level, being forthright, being respectful, and being thorough) are part of effective communication, which was validated by our findings in this study [74]. A poor physician-patient communication in cancer care negatively affects psychological well-being and patients' decisions and perceptions regarding treatments. This validates our findings of the impact of communication with physicians among newly diagnosed patients with cancer on their perception of technology use during visits and trust in physicians [75].

In addition, we found that the work system elements impact patients' workload, communication, and QOC perceptions. This correlates with the findings of other studies in which the environment design was shown to impact patients with cancer's perception of QOC. A recent review of evidence-based design also found that a conscious design adapted to patient needs had an impact on a decrease in infection spreading, length of stay, pharmacological needs, and perceived stress among patients [76]. Furthermore, symbolic objects found in the environment have been shown to impact patients' sense of self and well-being [76].

In addition, a recently published Cochrane review on environmental impact on health stressed the profound need for well-designed studies following intervention in health care environments [77]. This correlates with our finding that patients who liked the decoration of the hospital, the motivational signs, the colors, the cleanness, the organized processes, the lighting, and the care of the nurses were more satisfied with the QOC and felt less overwhelmed. These findings lead to the expectation that major considerations ought to be taken when designing health care environments to meet quality requirements while considering patients' needs and supporting patients' sense of control, autonomy, and independence.

Theoretical and Practical Implications of the Findings

In the previous section, we validated the preset hypotheses that correlate with the findings of the quantitative studies from the greater project. This framework can help inform patient-centered interventions that aim to provide newly diagnosed patients with cancer with the support needed and ensure their satisfaction with the QOC offered. More empathy and human bond links between physicians and patients should be considered as patients want to be treated in a more patient-centered way and to feel that they are not receiving the same care as everyone else in the same way.

Patients also want to have the chance to ask as many questions as possible and be given as many follow-up visits in the beginning as possible to receive comfort and reassurance that everything will be fine. Empowering workers (nurses and staff) to intervene in case of emergency would help patients trust the

XSL•FO
RenderX

health care organization. In addition, allowing patients to be accompanied by their family members would help them be emotionally comfortable. Another point to consider is to share a second screen with them in case a computer is used when the physician is communicating with them to comfort them regarding what the physicians are doing when they are not talking to them.

Pausing in the middle of the discussion to do other tasks would result in losing the patients' attention. Physicians should consider continuous communication where they pay as much attention as possible to the patients in a friendly way and where they listen to their concerns without rushing them even if there is a time limit as the time given can influence their decision-making process importantly.

This study's findings can also inform the organization's design. It should be considered that patients cannot move a lot between the laboratories and the visit rooms and it would be easier to assign them to rooms that are close to each other to minimize their physical effort during the visits. Better scheduling and allocation strategies should be considered to minimize the waiting time inside the hospital for each patient. Comforting colors, relaxing decoration, and motivational signs would help reassure patients while in the hospital. In addition, having any construction work when patients are coming in and out should be avoided as that can add more load to what they are already experiencing.

Limitations

Despite the useful insights garnered from this research, certain limitations must be addressed. First, the study's narrow geographic reach, which included only 1 cancer center, may limit the findings' generalizability to other cancer populations or health care settings. Patients' experiences and technology preferences in this facility may not represent those in other cancer centers or varied communities with different demographics or cultural backgrounds. Second, selection bias is possible as patients who chose to participate in the interviews may have different characteristics or opinions from those who declined or were unavailable. This may introduce bias in the findings and reduce the study's external validity. Furthermore, interviewing patients within a few visits following their initial diagnosis may not completely capture the dynamic character of their technological choices, which may change over time as patients adjust to their diagnosis and treatment. The reliance on patients' recollection of their technology preferences at this early time point may also be subject to recall bias. Furthermore, contextual factors particular to the cancer center where the research was conducted, such as local health care policies and the availability and accessibility of technology, may not be applicable or may vary in different contexts. Finally, social desirability bias and interviewer prejudice throughout the data collection process may have an impact on the data's authenticity and veracity. Despite these limitations, the findings of this study provide valuable insights into the technology preferences of newly diagnosed patients with cancer, and additional research with larger and more diverse samples, longer follow-up periods, and considerations of contextual factors is required to strengthen the findings' generalizability and validity.

Conclusions

In this study, we suggested a framework called Effectiveness of Patient-Centered Cancer Care and tested its validity in cancer visits to support PCC among newly diagnosed patients with cancer using qualitative data. We found that workload and patient-centered communication impact QOC and that the work system elements impact the patient-centeredness (workload and communication) and QOC (trust in physicians, satisfaction with care, and perception of technology use). To improve patients' experiences in the first visits after diagnosis, more interest needs to be given to the design of the organization, the processes that the patients have to go through, and the collaboration among the different actors and providers. This study's findings can also inform the organization's design. It should be considered that patients cannot move a lot between the laboratories and the visit rooms and it would be easier to assign them to rooms that are close to each other to minimize their physical effort during the visits. Better scheduling and allocation strategies should be considered to minimize the waiting time inside the hospital for each patient. Comforting colors, relaxing decoration, and motivational signs would help reassure patients while in the hospital. In addition, having any construction work when patients are coming in and out should be avoided as that can add more load to what they are already experiencing.

Acknowledgments

The research reported in this publication was supported by the National Institute of Nursing Research of the National Institutes of Health under award 1R15NR018965-01. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health. The authors also thank Dr Scott Rowley, Mariska Kop, Alene Mathurin, Farah Elkourdi, and Trina Mcclendon for their support in the data collection process.

Data Availability

The data sets generated during and analyzed during this study are available from the corresponding author on reasonable request.

Conflicts of Interest

None declared.

Multimedia Appendix 1

XSL•FO
RenderX

COREQ (Consolidated Criteria for Reporting Qualitative Research) checklist.
[DOCX File , 19 KB - humanfactors_v11i1e53053_app1.docx]

References

1. American Geriatrics Society Expert Panel on Person-Centered Care. Person-centered care: a definition and essential elements. J Am Geriatr Soc 2016 Jan;64(1):15-18. [doi: 10.1111/jgs.13866] [Medline: 26626262]
2. Li J, Porock D. Resident outcomes of person-centered care in long-term care: a narrative review of interventional research. Int J Nurs Stud 2014 Oct;51(10):1395-1415. [doi: 10.1016/j.ijnurstu.2014.04.003] [Medline: 24815772]
3. Elkefi S, Asan O. Perceived patient workload and its impact on outcomes during new cancer patient visits: analysis of a convenience sample. JMIR Hum Factors 2023 Aug 18;10:e49490 [FREE Full text] [doi: 10.2196/49490] [Medline: 37594798]
4. Elkefi S, Asan O. The impact of patient-centered care on cancer patients' QOC, self-efficacy, and trust towards doctors: analysis of a national survey. J Patient Exp 2023 Jan 17;10:23743735231151533 [FREE Full text] [doi: 10.1177/23743735231151533] [Medline: 36698621]
5. Mead N, Bower P. Measuring patient-centredness: a comparison of three observation-based instruments. Patient Educ Couns 2000 Jan;39(1):71-80. [doi: 10.1016/s0738-3991(99)00092-0] [Medline: 11013549]
6. Balint E. The possibilities of patient-centered medicine. J R Coll Gen Pract 1969 May;17(82):269-276 [FREE Full text] [Medline: 5770926]
7. Levenstein JH, McCracken EC, McWhinney IR, Stewart MA, Brown JB. The patient-centred clinical method. 1. A model for the doctor-patient interaction in family medicine. Fam Pract 1986 Mar;3(1):24-30. [doi: 10.1093/fampra/3.1.24] [Medline: 3956899]
8. Delbanco T, Berwick DM, Boufford JI, Edgman-Levitan S, Ollenschläger G, Plamping D, et al. Healthcare in a land called PeoplePower: nothing about me without me. Health Expect 2001 Sep 20;4(3):144-150 [FREE Full text] [doi: 10.1046/j.1369-6513.2001.00145.x] [Medline: 11493320]
9. National Academy of Sciences. The National Academies Collection: Reports Funded by National Institutes of Health. Washington, DC: National Academies Press; 2008.
10. Constand MK, MacDermid JC, Dal Bello-Haas D, Law M. Scoping review of patient-centered care approaches in healthcare. BMC Health Serv Res 2014 Jun 19;14(1):271 [FREE Full text] [doi: 10.1186/1472-6963-14-271] [Medline: 24947822]
11. Framework on integrated, people-centred health services. World Health Organization. 2016. URL: https://apps.who.int/gb/ebwha/pdf_files/WHA69/A69_39-en.pdf [accessed 2024-04-05]
12. Miller KL. Patient centered care: a path to better health outcomes through engagement and activation. NeuroRehabilitation 2016 Oct 14;39(4):465-470. [doi: 10.3233/NRE-161378] [Medline: 27689606]
13. Tevaarwerk AJ, Klemp JR, van Londen GJ, Hesse BW, Sesto ME. Moving beyond static survivorship care plans: a systems engineering approach to population health management for cancer survivors. Cancer 2018 Nov 15;124(22):4292-4300 [FREE Full text] [doi: 10.1002/cncr.31546] [Medline: 30277575]
14. Balbale SN, Turcios S, LaVela SL. Health care employee perceptions of patient-centered care. Qual Health Res 2015 Mar 01;25(3):417-425 [FREE Full text] [doi: 10.1177/1049732314553011] [Medline: 25274626]
15. Bergeson SC, Dean JD. A systems approach to patient-centered care. JAMA 2006 Dec 20;296(23):2848-2851. [doi: 10.1001/jama.296.23.2848] [Medline: 17179462]
16. Carayon P, Wetterneck TB, Rivera-Rodriguez AJ, Hundt AS, Hoonakker P, Holden R, et al. Human factors systems approach to healthcare quality and patient safety. Appl Ergon 2014 Jan;45(1):14-25 [FREE Full text] [doi: 10.1016/j.apergo.2013.04.023] [Medline: 23845724]
17. Schunk DH. Social cognitive theory. In: Harris KR, Graham S, Urdan T, McCormick CB, Sinatra GM, Sweller J, editors. APA Educational Psychology Handbook, Vol. 1. Theories, Constructs, and Critical Issues. New York, NY: American Psychological Association; 2012.
18. Epstein R, Street R. Patient-centered communication. In: Wittenberg E, Ferrell BR, Goldsmith J, Smith T, Glajchen M, Handzo RG, editors. Textbook of Palliative Care Communication. Oxford, UK: Oxford University Press; 2007:17-38.
19. Venkatesh V, Davis FD. A theoretical extension of the technology acceptance model: four longitudinal field studies. Manag Sci 2000 Feb;46(2):186-204. [doi: 10.1287/mnsc.46.2.186.11926]
20. Marangunić N, Granić A. Technology acceptance model: a literature review from 1986 to 2013. Univ Access Inf Soc 2014 Feb 16;14(1):81-95. [doi: 10.1007/s10209-014-0348-1]
21. Carayon P, Schoofs Hundt A, Karsh BT, Gurses AP, Alvarado CJ, Smith M, et al. Work system design for patient safety: the SEIPS model. Qual Saf Health Care 2006 Dec;15 Suppl 1(Suppl 1):i50-i58 [FREE Full text] [doi: 10.1136/qshc.2005.015842] [Medline: 17142610]
22. Carayon P, Wooldridge A, Hoonakker P, Hundt AS, Kelly MM. SEIPS 3.0: Human-centered design of the patient journey for patient safety. Appl Ergon 2020 Apr;84:103033 [FREE Full text] [doi: 10.1016/j.apergo.2019.103033] [Medline: 31987516]
23. Li J, Luo X, Cao Q, Lin Y, Xu Y, Li Q. Communication needs of cancer patients and/or caregivers: a critical literature review. J Oncol 2020;2020:7432849 [FREE Full text] [doi: 10.1155/2020/7432849] [Medline: 32454826]

XSL•FO
RenderX

24. Cheng AC, Levy MA. Measures of treatment workload for patients with breast cancer. JCO Clin Cancer Inform 2019 Feb;3:1-10 [FREE Full text] [doi: 10.1200/CCI.18.00122] [Medline: 30715929]
25. Holden RJ, Cornet VP, Valdez RS. Patient ergonomics: 10-year mapping review of patient-centered human factors. Appl Ergon 2020 Jan;82:102972 [FREE Full text] [doi: 10.1016/j.apergo.2019.102972] [Medline: 31654954]
26. Holden RJ, Carayon P, Gurses AP, Hoonakker P, Hundt AS, Ozok AA, et al. SEIPS 2.0: a human factors framework for studying and improving the work of healthcare professionals and patients. Ergonomics 2013 Nov;56(11):1669-1686 [FREE Full text] [doi: 10.1080/00140139.2013.838643] [Medline: 24088063]
27. Norris B, West J, Anderson O, Davey G, Brodie A. Taking ergonomics to the bedside--a multi-disciplinary approach to designing safer healthcare. Appl Ergon 2014 May;45(3):629-638. [doi: 10.1016/j.apergo.2013.09.004] [Medline: 24135560]
28. Zimmerman ME. Task load. In: Kreutzer JS, DeLuca J, Caplan B, editors. Encyclopedia of Clinical Neuropsychology. New York, NY: Springer; 2011:2469-2470.
29. Korbach A, Brünken R, Park B. Differentiating different types of cognitive load: a comparison of different measures. Educ Psychol Rev 2017 Mar 31;30(2):503-529. [doi: 10.1007/s10648-017-9404-8]
30. Miller S. Workload measures: national advanced driving simulator. The University of Iowa. 2001. URL: https://www.nads-sc.uiowa.edu/publicationStorage/200501251347060.N01-006.pdf [accessed 2024-04-05]
31. LeBlanc TW, Back AL, Danis M, Abernethy AP. Electronic Health Records (EHRs) in the oncology clinic: how clinician interaction with EHRs can improve communication with the patient. J Oncol Pract 2014 Sep;10(5):317-321 [FREE Full text] [doi: 10.1200/JOP.2014.001385] [Medline: 25027025]
32. Hack TF, Degner LF, Parker PA, SCRN Communication Team. The communication goals and needs of cancer patients: a review. Psychooncology 2005 Oct;14(10):831-846. [doi: 10.1002/pon.949] [Medline: 16200519]
33. Eggly S, Penner LA, Hagiwara N, Gonzalez R, Harper FW, Heath EI, et al. Patient, companion, and oncologist agreement regarding information discussed during triadic oncology clinical interactions. Psychooncology 2013 Mar 15;22(3):637-645 [FREE Full text] [doi: 10.1002/pon.3045] [Medline: 22337320]
34. Wagner EH, Aiello Bowles EJ, Greene SM, Tuzzio L, Wiese CJ, Kirlin B, et al. The quality of cancer patient experience: perspectives of patients, family members, providers and experts. Qual Saf Health Care 2010 Dec 02;19(6):484-489. [doi: 10.1136/qshc.2010.042374] [Medline: 21127109]
35. Klasnja P, Hartzler AC, Unruh KT, Pratt W. Blowing in the wind: unanchored patient information work during cancer care. Proc SIGCHI Conf Hum Factor Comput Syst 2010 Apr;2010:193-202 [FREE Full text] [doi: 10.1145/1753326.1753355] [Medline: 20697453]
36. Makoul G, Krupat E, Chang CH. Measuring patient views of physician communication skills: development and testing of the Communication Assessment Tool. Patient Educ Couns 2007 Aug;67(3):333-342 [FREE Full text] [doi: 10.1016/j.pec.2007.05.005] [Medline: 17574367]
37. Batalden P, Leach D, Swing S, Dreyfus H, Dreyfus S. General competencies and accreditation in graduate medical education. Health Aff (Millwood) 2002 Sep;21(5):103-111. [doi: 10.1377/hlthaff.21.5.103] [Medline: 12224871]
38. Horowitz SD. Evaluation of clinical competencies: basic certification, subspecialty certification, and recertification. Am J Phys Med Rehabil 2000;79(5):478-480. [doi: 10.1097/00002060-200009000-00016] [Medline: 10994892]
39. Duffy FD, Gordon GH, Whelan G, Cole-Kelly K, Frankel R, Buffone N, Participants in the American Academy on Physician and Patient's Conference on Education and Evaluation of Competence in Communication and Interpersonal Skills. Assessing competence in communication and interpersonal skills: the Kalamazoo II report. Acad Med 2004 Jun;79(6):495-507. [doi: 10.1097/00001888-200406000-00002] [Medline: 15165967]
40. Massie MJ, Holland JC. The cancer patient with pain: psychiatric complications and their management. J Pain Symptom Manage 1992 Feb;7(2):99-109 [FREE Full text] [doi: 10.1016/0885-3924(92)90121-w] [Medline: 1573292]
41. Ong LM, Visser MR, Lammes FB, de Haes JC. Doctor-patient communication and cancer patients' quality of life and satisfaction. Patient Educ Couns 2000 Sep;41(2):145-156. [doi: 10.1016/s0738-3991(99)00108-1] [Medline: 12024540]
42. Stewart MA. What is a successful doctor-patient interview? A study of interactions and outcomes. Soc Sci Med 1984;19(2):167-175 [FREE Full text] [doi: 10.1016/0277-9536(84)90284-3] [Medline: 6474233]
43. Roter DL, Hall JA, Katz NR. Relations between physicians' behaviors and analogue patients' satisfaction, recall, and impressions. Med Care 1987 May;25(5):437-451. [doi: 10.1097/00005650-198705000-00007] [Medline: 3695654]
44. Kaplan SH, Greenfield S, Ware Jr JE. Assessing the effects of physician-patient interactions on the outcomes of chronic disease. Med Care 1989 Mar;27(3 Suppl):S110-S127. [doi: 10.1097/00005650-198903001-00010] [Medline: 2646486]
45. Henbest RJ, Stewart M. Patient-centredness in the consultation. 2: does it really make a difference? Fam Pract 1990 Mar;7(1):28-33. [doi: 10.1093/fampra/7.1.28] [Medline: 2318368]
46. Buntin MB, Burke MF, Hoaglin MC, Blumenthal D. The benefits of health information technology: a review of the recent literature shows predominantly positive results. Health Aff (Millwood) 2011 Mar;30(3):464-471. [doi: 10.1377/hlthaff.2011.0178] [Medline: 21383365]
47. Jones SS, Rudin RS, Perry T, Shekelle PG. Health information technology: an updated systematic review with a focus on meaningful use. Ann Intern Med 2014 Jan 07;160(1):48-54 [FREE Full text] [doi: 10.7326/M13-1531] [Medline: 24573664]

XSL•FO
RenderX

48. Sligo J, Gauld R, Roberts V, Villa L. A literature review for large-scale health information system project planning, implementation and evaluation. Int J Med Inform 2017 Jan;97:86-97. [doi: 10.1016/j.ijmedinf.2016.09.007] [Medline: 27919399]
49. Epstein RM, Street RL. The values and value of patient-centered care. Ann Fam Med 2011 Mar 14;9(2):100-103 [FREE Full text] [doi: 10.1370/afm.1239] [Medline: 21403134]
50. ElKefi S, Asan O. How technology impacts communication between cancer patients and their health care providers: a systematic literature review. Int J Med Inform 2021 May;149:104430 [FREE Full text] [doi: 10.1016/j.ijmedinf.2021.104430] [Medline: 33684711]
51. Hardiker NR, Grant MJ. Factors that influence public engagement with eHealth: a literature review. Int J Med Inform 2011 Jan;80(1):1-12 [FREE Full text] [doi: 10.1016/j.ijmedinf.2010.10.017] [Medline: 21112244]
52. Meskó B, Drobni Z, Bényei É, Gergely B, Győrffy Z. Digital health is a cultural transformation of traditional healthcare. Mhealth 2017 Sep;3:38 [FREE Full text] [doi: 10.21037/mhealth.2017.08.07] [Medline: 29184890]
53. Asan O. Providers' perceived facilitators and barriers to EHR screen sharing in outpatient settings. Appl Ergon 2017 Jan;58:301-307. [doi: 10.1016/j.apergo.2016.07.005] [Medline: 27633226]
54. Aapro M, Bossi P, Dasari A, Fallowfield L, Gascón P, Geller M, et al. Digital health for optimal supportive care in oncology: benefits, limits, and future perspectives. Support Care Cancer 2020 Oct 12;28(10):4589-4612 [FREE Full text] [doi: 10.1007/s00520-020-05539-1] [Medline: 32533435]
55. Mazur LM, Mosaly PR, Moore C, Marks L. Association of the usability of electronic health records with cognitive workload and performance levels among physicians. JAMA Netw Open 2019 Apr 05;2(4):e191709 [FREE Full text] [doi: 10.1001/jamanetworkopen.2019.1709] [Medline: 30951160]
56. Mazor KM, Beard RL, Alexander GL, Arora NK, Firneno C, Gaglio B, et al. Patients' and family members' views on patient-centered communication during cancer care. Psychooncology 2013 Nov;22(11):2487-2495 [FREE Full text] [doi: 10.1002/pon.3317] [Medline: 23780672]
57. Hillen MA, de Haes HC, Smets EM. Cancer patients' trust in their physician-a review. Psychooncology 2011 Mar 23;20(3):227-241. [doi: 10.1002/pon.1745] [Medline: 20878840]
58. Cao W, Qi X, Yao T, Han X, Feng X. How doctors communicate the initial diagnosis of cancer matters: cancer disclosure and its relationship with Patients' hope and trust. Psychooncology 2017 May 17;26(5):640-648. [doi: 10.1002/pon.4063] [Medline: 26776310]
59. Urden LD. Patient satisfaction measurement: current issues and implications. Lippincotts Case Manag 2002;7(5):194-200. [doi: 10.1097/00129234-200209000-00006] [Medline: 12394558]
60. Mallinger JB, Griggs JJ, Shields CG. Patient-centered care and breast cancer survivors' satisfaction with information. Patient Educ Couns 2005 Jun;57(3):342-349. [doi: 10.1016/j.pec.2004.09.009] [Medline: 15893218]
61. Venetis MK, Robinson JD, Turkiewicz KL, Allen M. An evidence base for patient-centered cancer care: a meta-analysis of studies of observed communication between cancer specialists and their patients. Patient Educ Couns 2009 Dec;77(3):379-383. [doi: 10.1016/j.pec.2009.09.015] [Medline: 19836920]
62. Mazor KM, Roblin DW, Greene SM, Lemay CA, Firneno CL, Calvi J, et al. Toward patient-centered cancer care: patient perceptions of problematic events, impact, and response. J Clin Oncol 2012 May 20;30(15):1784-1790 [FREE Full text] [doi: 10.1200/JCO.2011.38.1384] [Medline: 22508828]
63. Peppercorn JM, Smith TJ, Helft PR, Debono DJ, Berry SR, Wollins DS, American Society of Clinical Oncology. American society of clinical oncology statement: toward individualized care for patients with advanced cancer. J Clin Oncol 2011 Feb 20;29(6):755-760. [doi: 10.1200/JCO.2010.33.1744] [Medline: 21263086]
64. Gallagher TH. A 62-year-old woman with skin cancer who experienced wrong-site surgery: review of medical error. JAMA 2009 Aug 12;302(6):669-677. [doi: 10.1001/jama.2009.1011] [Medline: 19584321]
65. Surbone A. Telling the truth to patients with cancer: what is the truth? Lancet Oncol 2006 Nov;7(11):944-950. [doi: 10.1016/s1470-2045(06)70941-x]
66. Ayanian JZ, Zaslavsky AM, Guadagnoli E, Fuchs CS, Yost KJ, Creech CM, et al. Patients' perceptions of quality of care for colorectal cancer by race, ethnicity, and language. J Clin Oncol 2005 Sep 20;23(27):6576-6586. [doi: 10.1200/JCO.2005.06.102] [Medline: 16116149]
67. Elkefi S. Evaluation of patient-centeredness in cancer care: a mixed methods study with a human factors approach. Stevens Institute of Technology. URL: https://www.proquest.com/openview/620d49593ddb80f77648a87fc6c0e905/1?pq-origsite=gscholar&cbl=18750&diss=y [accessed 2024-04-05]
68. Charmaz K. Constructing Grounded Theory: A Practical Guide through Qualitative Analysis. New York, NY: Sage Publications; 2006.
69. Nelson J. Using conceptual depth criteria: addressing the challenge of reaching saturation in qualitative research. Qual Res 2016 Dec 14;17(5):554-570. [doi: 10.1177/1468794116679873]
70. Lewandowska A, Rudzki G, Lewandowski T, Rudzki S. The problems and needs of patients diagnosed with cancer and their caregivers. Int J Environ Res Public Health 2020 Dec 24;18(1):87 [FREE Full text] [doi: 10.3390/ijerph18010087] [Medline: 33374440]

XSL•FO
RenderX

71. Page AE, Adler NE. Cancer Care for the Whole Patient Meeting: Psychosocial Health Needs. Washington, DC: National Academies Press; 2021.
72. Avancini A, Pala V, Trestini I, Tregnago D, Mariani L, Sieri S, et al. Exercise levels and preferences in cancer patients: a cross-sectional study. Int J Environ Res Public Health 2020 Jul 24;17(15):5351 [FREE Full text] [doi: 10.3390/ijerph17155351] [Medline: 32722265]
73. Bager L, Elsbernd A, Nissen A, Daugaard G, Pappot H. Danish translation and pilot testing of the European Organization for Research and Treatment of Cancer QLQ-TC 26 (EORTC QLQ-TC26) questionnaire to assess health-related quality of life in patients with testicular cancer. Health Qual Life Outcomes 2018 Jun 18;16(1):128 [FREE Full text] [doi: 10.1186/s12955-018-0954-3] [Medline: 29914500]
74. Grassi L, Caruso R, Costantini A. Communication with patients suffering from serious physical illness. Adv Psychosom Med 2015;34:10-23. [doi: 10.1159/000369050] [Medline: 25832510]
75. Lehmann C, Koch U, Mehnert A. [Impact of the doctor-patient-communication on distress and utilization of psychosocial services among cancer patients. A review of the current literature]. Psychother Psychosom Med Psychol 2009 Jul;59(7):e3-27. [doi: 10.1055/s-2008-1067443] [Medline: 18666041]
76. Browall M, Koinberg I, Falk H, Wijk H. Patients' experience of important factors in the healthcare environment in oncology care. Int J Qual Stud Health Well-being 2013 Aug 06;8(1):20870 [FREE Full text] [doi: 10.3402/qhw.v8i0.20870] [Medline: 23924604]
77. Drahota A, Ward D, Mackenzie H, Stores R, Higgins B, Gal D, et al. Sensory environment on health-related outcomes of hospital patients. Cochrane Database Syst Rev 2012 Mar 14;2012(3):CD005315. [doi: 10.1002/14651858.CD005315.pub2] [Medline: 22419308]

Abbreviations

COREQ: Consolidated Criteria for Reporting Qualitative Research
HIT: health IT
PCC: patient-centered care
QOC: quality of care
SEIPS: Systems Engineering Initiative for Patient Safety

Edited by A Choudhury; submitted 24.09.23; peer-reviewed by Y Luo, F Albalwy; comments to author 08.02.24; revised version received 06.03.24; accepted 15.03.24; published 24.04.24.

Please cite as:
Elkefi S, Asan O
Validating the Effectiveness of the Patient-Centered Cancer Care Framework by Assessing the Impact of Work System Factors on Patient-Centered Care and Quality of Care: Interview Study With Newly Diagnosed Cancer Patients
JMIR Hum Factors 2024;11:e53053
URL: https://humanfactors.jmir.org/2024/1/e53053
doi:10.2196/53053
PMID:38656776

©Safa Elkefi, Onur Asan. Originally published in JMIR Human Factors (https://humanfactors.jmir.org), 24.04.2024. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on https://humanfactors.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Development of a Real-Time Dashboard for Overdose Touchpoints: User-Centered Design Approach

Amey Salvi[1], MS; Logan A Gillenwater[2], BS; Brandon P Cockrum[3], MFA; Sarah E Wiehe[2,3], MPH, MD; Kaitlyn Christian[4], MPA; John Cayton[4], MAEd, MS; Timothy Bailey[4], MS; Katherine Schwartz[2], MPA, JD; Allyson L Dir[5], PhD; Bradley Ray[6], PhD; Matthew C Aalsma[2], PhD; Khairi Reda[1], PhD

[1]School of Informatics, Computing, and Engineering, Indiana University Indianapolis, Indianapolis, IN, United States
[2]Children's Health Services Research, Department of Pediatrics, Indiana University School of Medicine, Indianapolis, IN, United States
[3]Indiana Clinical and Translational Science Institute, Indianapolis, IN, United States
[4]Indiana Management Performance Hub, Indianapolis, IN, United States
[5]Department of Psychiatry, Indiana University School of Medicine, Indianapolis, IN, United States
[6]RTI International, Research Triangle Park, NC, United States

Corresponding Author:
Khairi Reda, PhD
School of Informatics, Computing, and Engineering
Indiana University Indianapolis
535 W Michigan St
Indianapolis, IN, 46202
United States
Phone: 1 317 274 5788
Email: redak@iu.edu

Abstract

Background: Overdose Fatality Review (OFR) is an important public health tool for shaping overdose prevention strategies in communities. However, OFR teams review only a few cases at a time, which typically represent a small fraction of the total fatalities in their jurisdiction. Such limited review could result in a partial understanding of local overdose patterns, leading to policy recommendations that do not fully address the broader community needs.

Objective: This study explored the potential to enhance conventional OFRs with a data dashboard, incorporating visualizations of touchpoints—events that precede overdoses—to highlight prevention opportunities.

Methods: We conducted 2 focus groups and a survey of OFR experts to characterize their information needs and design a real-time dashboard that tracks and measures decedents' past interactions with services in Indiana. Experts (N=27) were engaged, yielding insights on essential data features to incorporate and providing feedback to guide the development of visualizations.

Results: The findings highlighted the importance of showing decedents' interactions with health services (emergency medical services) and the justice system (incarcerations). Emphasis was also placed on maintaining decedent anonymity, particularly in small communities, and the need for training OFR members in data interpretation. The developed dashboard summarizes key touchpoint metrics, including prevalence, interaction frequency, and time intervals between touchpoints and overdoses, with data viewable at the county and state levels. In an initial evaluation, the dashboard was well received for its comprehensive data coverage and its potential for enhancing OFR recommendations and case selection.

Conclusions: The Indiana touchpoints dashboard is the first to display real-time visualizations that link administrative and overdose mortality data across the state. This resource equips local health officials and OFRs with timely, quantitative, and spatiotemporal insights into overdose risk factors in their communities, facilitating data-driven interventions and policy changes. However, fully integrating the dashboard into OFR practices will likely require training teams in data interpretation and decision-making.

(JMIR Hum Factors 2024;11:e57239) doi:10.2196/57239

XSL•FO
RenderX

KEYWORDS

overdose prevention; dashboards; fatality review; data integration; visualizations; visualization; dashboard; fatality; death; overdose; overdoses; overdosing; prevention; develop; development; design; interview; interviews; focus group; focus groups; touchpoints; touchpoint; substance abuse; drug abuse

Introduction

Background

The escalating drug overdose epidemic in the United States continues to pose a major public health challenge. Previous research has identified general risk factors that are linked to increased overdose rates [1-3], including unstable housing [4,5], recent release from incarceration [6,7], and frequent visits to the emergency department (ED) [8-11]. However, overdose risk factors exhibit considerable variation across communities and are influenced heavily by geographic and demographic disparities, particularly in access to health care and prevention services [9,12]. Moreover, the evolving nature of the epidemic has led to shifting risk profiles among different subpopulations [13]. These disparities underscore the need for timely and data-driven interventions that are tailored to the specific needs and challenges of local communities.

One mechanism for implementing targeted, community-specific interventions is through local Overdose Fatality Reviews (OFRs). Modeled after child fatality reviews [14,15], OFR teams comprise reviewers from multiple agencies who conduct collaborative, in-depth reviews of case files for individuals who have died of overdose [16,17]. Through these detailed case reviews, OFRs identify service gaps and recommend strategies to prevent future overdoses in their communities. The use of OFRs has gained momentum, with teams operating across various US localities [18]. However, current OFR practices primarily focus on reviewing only a handful of cases, typically 2 to 5 monthly or quarterly [19]. These cases typically represent a small fraction of the total fatalities occurring in their jurisdiction. While informative, the emphasis on a few individual cases could skew the review process, leading to OFRs making recommendations that do not fully address broader overdose trends.

As local governments continue to collect data on overdose events, there is an opportunity to leverage these data to enhance the OFR process. Previous research demonstrates the value of linking administrative data sets routinely collected by state governments (eg, calls to emergency services and incarceration records) with overdose mortality data [20-24]. For example, cross-referencing the records of decedents who experienced overdoses from across various data sets allows for uncovering their "touchpoints"—interactions with health and social services and other local systems they had before their overdose. When brought to light, touchpoints offer key opportunities to engage at-risk individuals and connect them with prevention services and treatments [25-27]. Analyses to identify touchpoints have so far been performed manually by researchers. However, the process is amenable to automation, enabling continuous assessment of touchpoint characteristics. The results can then be communicated in real time to local OFRs through a dashboard, providing review teams with up-to-date, quantitative information on the trajectories of decedents in their communities.

Dashboards have proven invaluable in public health settings [28,29] owing to their ability to visually summarize key metrics and statistics [30,31], thereby aiding surveillance and fostering evidence-based responses to emerging health threats [32,33]. Furthermore, dashboards are conducive to collaborative sense making among multiple individuals [34-36]. This feature makes them particularly suited to fatality review meetings, which are designed to be collaborative and deliberative in nature. Numerous dashboards have been developed to visualize drug overdose–related data [37-39]. However, existing solutions are primarily intended to surveil the level and distribution of overdoses as opposed to understanding events that *precede* them. Few of the earlier dashboards showcase touchpoints at the local level or update data in real time, making them less suited for understanding system-level gaps or for deriving prevention-oriented insights.

Aims

This study presents findings from human-centered research, design, development, and initial evaluation of a dashboard aimed at supporting OFR teams by visualizing overdose touchpoint statistics. The objective was to provide county-level OFR teams with timely and actionable data on events that consistently *precede* fatal overdoses in their communities. In doing so, we aimed to illuminate additional opportunities for interventions at the population level beyond what can be gleaned from individual fatality case reviews. The goal was to increase the chance of successful targeting and implementation of OFR recommendations. This stands to improve overdose prevention and reduce the number of preventable deaths.

Methods

Overview

To design a dashboard suitable for the needs of OFRs, we adopted a user-centered design framework [40,41] drawing on participatory methods to engage stakeholders in the process [42,43]. Specifically, we conducted focus groups with a panel of OFR experts to elicit perspectives on requirements and data needs, envision design possibilities, and document potential challenges. The elicited requirements were then used to develop exploratory visualizations of touchpoints data. The initial visualizations were further refined based on feedback from the expert panel. Subsequently, the revised visualizations were used to develop a web-based dashboard that is hosted by the Indiana state government.

Study Setting and Data Sources

We partnered with the state government of Indiana to prototype and develop the sought touchpoints dashboard. Indiana has a nationally recognized role in organizing and convening OFRs, with 28 active review teams organized at the county level and

XSL•FO
RenderX

supported by the Indiana Department of Health. Similar to many other states, Indiana maintains a comprehensive and up-to-date database of fatal overdoses. This database includes all suspected accidental poisonings (coded as X40-X44), intentional poisonings (X60-X64), assaults by drug (X85), and cases of undetermined intent (Y10-Y14) that occurred among Indiana residents. In addition to overdose data, the state maintains administrative data sets from various agencies, including incarceration records, emergency and medical service use, and prescription dispensation. Importantly, these administrative data sets are linkable to the overdose mortality records. The Indiana Management Performance Hub (MPH), a state-level agency, serves as a central repository for these data sets, which are gathered from the corresponding agencies.

To identify events that precede drug-related fatalities, overdose cases are linked to administrative data sets at the individual level. This linking procedure is performed by the MPH using a probabilistic matching algorithm that considers identifiers such as the decedent's name, date of birth, and social security number, among others. This process allows for the reconstruction of past interactions with various touchpoints for each identifiable decedent. Subsequently, deidentified statistics about these interactions are pushed to the dashboard for visualization. This linkage process is performed weekly, enabling (near) real-time updates of the visualizations.

User-Centered Design Process

To inform the design of the dashboard, we conducted 2 focus groups with a panel of OFR experts. We recruited participants via email, inviting experienced OFR practitioners and early developers from across the United States. Our goal in these focus groups was to understand OFR information needs and leverage the panel's experience in conceptualizing, co-designing, and refining visualizations. The focus groups took place virtually using Zoom videoconferencing software (Zoom Video Communications). A virtual whiteboard was used to place and arrange "Post-it"–style notes. Participating experts were recruited from the same pool, with later focus groups involving fewer participants to allow for convergence and facilitate more in-depth feedback. The focus groups were video recorded, transcribed, and analyzed using thematic analysis techniques [44].

The first focus group sought to uncover data access barriers and needs for OFR teams. A total of 13 experts participated in the discussion. Participants were first prompted to share challenges and "pain points" regarding access to data. In a second activity, participants were divided into 2 breakout groups to identify key data attributes essential for review teams. They also gave high-level design parameters for the dashboard. Finally, participants reflected on their hopes and concerns for the dashboard's integration into OFR processes, emphasizing potential positive outcomes and addressing apprehensions.

On the basis of the findings of the initial focus group, we created a series of 6 initial visualizations that illustrate overdose touchpoints using a static snapshot of the MPH-linked data set described previously. These initial visualizations served as the foundation for a second focus group with the participation of 6 experts. During this session, a facilitator presented each of the 6 visualizations and prompted participants for feedback. Specifically, participants were asked to evaluate the ease of understanding of these visualizations and their potential usefulness in the OFR process. We sought additional input by conducting a survey of 5 experts. The survey presented the same initial visualizations and requested open-ended comments on their intuitiveness and utility. Insights gathered from the survey along with feedback obtained during the second focus group were used to refine the visualizations and develop an interactive dashboard.

Dashboard Evaluation

To obtain feedback on the final dashboard, we conducted an initial assessment with 3 OFR experts. Participants were asked to perform a series of data extraction tasks (eg, identifying the touchpoint with the highest prevalence). In addition, they were prompted to make recommendations based on the observed touchpoint patterns, simulating the use of the dashboard within a typical OFR meeting.

Ethical Considerations

This human-centered research was reviewed and approved by the Indiana University institutional review board (approval 17809). Participants received an information sheet explaining the study goals and procedures before agreeing to take part. The analysis of state mortality and administrative data sets, while not considered human participant research, followed state legal and ethical procedures. The dashboard displays only aggregate, population-level visualizations. No individual records are released or displayed to preserve anonymity. Furthermore, special care was taken to minimize the risk of reidentification by withholding actual event counts and substituting with percentages. Participants received a US $100 gift card as compensation.

Results

Overview

Participants highlighted barriers faced by OFRs in accessing and interpreting data within the context of fatality reviews. They also provided insights on what data attributes and features would be most useful for OFRs to look at. We report these findings and discuss how we incorporated them to create a real-time dashboard for visualizing overdose touchpoints.

Barriers to Accessing and Using Data

Data Accessibility

Several participants highlighted the lack of access to data as one of the major barriers in fatality reviews. Some of these barriers stem from challenges in sharing available data due to legal restrictions, data security, and privacy concerns:

> *Asking our state offices for data would result in, "Sorry, we can't share on the state level." There [needs to] be intergovernmental agreements between state police or our mental health or our human services or our health department.* [P2]

> *[Gaining access] is always an issue, and especially without laws that allow for the OFRs to get this. I*

XSL•FO
RenderX

> *know we had a lot of laws related to the child death review teams that I worked with that allowed us access to data, but it wasn't always the same for other death review teams.* [P11]

While recognizing existing regulatory and logistical obstacles, participants anticipated that increasing data access could empower OFRs to make more informed decisions:

> *We're trying to drive positive change that could maybe be implemented statewide, and they just give us a little bit. It [data] would give us the power to make better decisions.* [P2]

In addition to data access, the quality and accuracy of the data were also brought up as a prominent issue for OFRs, especially because of acknowledged variations in how data are coded and measured across different organizations. For example, 1 participant cited different standards for classifying services, noting that such inconsistencies could lead to misinterpretation:

> *When it's really law enforcement heavy, they're not understanding the public health ramifications of criminal justice involvement. It affects the lens from which data's being collected. So, when I go through the qualitative data...we've got people identifying jail substance use services as harm reduction, [and] you end up collecting some inaccurate data, which then misinforms the big picture.* [P7]

Influx of Case-Specific Data

While obtaining population-level data in certain arenas proved challenging, another concern was the vast amount of case-specific data that OFRs must already contend with. Participants noted that review teams are increasingly tasked with handling large volumes of individual reports from multiple systems, which often need to be manually and qualitatively analyzed at considerable time and effort:

> *OFRs collect an enormous amount of data, but you really need a whole army of researchers to be able to analyze it, especially the qualitative data. When the teams are putting forth all of these recommendations, it's just so hard to go through all the information and make a meaningful plan of it.* [P7]

Extensive data on individual death circumstances (as opposed to population-level statistics) reflect a conventional OFR focus on in-depth reviews of a few strategically selected cases. However, with the sheer number of overdose fatalities, it becomes difficult for OFRs to ensure that the selected cases represent the broader overdose patterns and risk factors prevalent in their community. One participant put it as follows:

> *[My experience] is that they would just randomly pick cases and then do a really deep dive into those cases, but you have no way to actually ensure that those are representative...And so, my hope had been that we would have certain [data] fields that we could have someone enter, and then that would allow us to do really large-scale analysis over the course of multiple years...[This] would have allowed us to really have a good sense as it relates to a variety of factors, but there just wasn't capacity. So, then we're just picking cases that look good or meet some theme to be able to have a more robust conversation at any given meeting. But again, they're not necessarily representative and you don't end up having the whole picture.* [P18]

Key Data Types and Attributes for the Dashboard

Participants identified key data attributes that they deemed essential for inclusion in a dashboard. We divided these attributes into 3 categories: touchpoints, social determinants of overdose risk, and case-specific data.

Touchpoints

Touchpoints represent interactions with systems and services before overdose. Thus, they serve as opportunities to connect people who use drugs with additional prevention services and treatments, potentially mitigating the risk of future overdoses. A frequently recurring set of touchpoints identified by experts was interaction with the justice system. For instance, the duration between a decedent's overdose and their last incarceration or residential treatment was cited as particularly important:

> *Were they justice involved or not at any point, but also the average distance in time from their last incarceration...So, to see were they in that window of high risk. And same if they were in residential treatments as average number of days.* [P3]

> *Average days out from treatment and incarceration because I feel like those are solid spaces that action can be taken.* [P5]

Several participants pointed to interactions with justice systems broadly as key touchpoints. Agencies such as county sheriffs, local police departments, and child protective services were thought to play a crucial role in an individual's risk of overdose both positively and negatively:

> *Justice systems can either be a force of treatment or a barrier to treatment. I think that involvement is really important...the extent of involvement can be really helpful to inform the justice system and the legislative changes that could help.* [P4]

Participants noted that data on criminal justice touchpoints might reveal new prevention opportunities or support policy recommendations, such as facilitating continued treatment for institutionalized individuals:

> *...keep people engaged in treatment, [such that] we're not disrupting treatment by violating [ie, rearresting] people and incarcerating them...It's a fruitful area for policy change. Most of our policy changes and recommendations from our OFR have been in the justice space.* [P3]

In addition to justice systems, participants noted interactions with health and medical facilities as crucial touchpoints. This included visits to the ED and emergency medical services (EMS):

XSL•FO
RenderX

> *Do we have one [attribute] here [on] the last date of medical intervention? Maybe like an ED visit or anything like that?* [P4]

> *There's an ED and EMS interaction right at the center there.* [P5]

Overall, three primary touchpoint categories emerged: (1) encounters with the justice system, such as incarceration; (2) engagement with health services, including ED and EMS interactions; and (3) involvement with residential treatment services. These touchpoints were recognized by participants as crucial opportunities for understanding risk factors and implementing services to close treatment gaps. Importantly, participants emphasized the typical interval between these touchpoints and overdose events as a critical feature to emphasize in the dashboard.

Social Determinants of Health

A second set of data attributes identified pertained to the social condition of the individuals themselves, which could shed light on factors that contribute to elevated overdose risk. For example, one of these factors was demographics:

> *Basic demographic information like poverty level, education level, homelessness. Anything that would affect those social determinants of health.* [P4]

A second factor was individuals' access to harm reduction services, as the same participant noted:

> *I was going to add...access to harm reduction services. So, what an environmental scan of resources or access to naloxone, treatment centers, syringe service programs, all those different community level access points.* [P4]

A third factor was housing, encompassing the shelter system and housing agencies:

> *Access to housing. Or maybe it's access to shelter because it could be both. There's housing policy, but then there's also the shelter systems.* [P5]

A fourth factor was the availability of transportation, which, according to participants, could influence an individual's access to treatment and harm reduction services:

> *Transportation between places: how easy is it for someone to get from point A to point B? Even if there's a syringe service program down the street, can they get to it? That kind of thing.* [P5]

Finally, participants also identified upstream social determinants such adverse childhood experiences as potentially relevant factors in assessing overdose risk:

> *...and some of that I think would fall under ACEs too because even if they're an adult, finding out if they were involved in that system as a child, trying to make some of those associations maybe.* [P5]

Case-Specific Data

Alongside touchpoints and social determinants of health, participants cited certain case-specific data, including toxicology reports, interviews with next of kin, and the decedent's circumstances at the time of death (eg, their position and whether they were alone). While these attributes are relevant to reviewing individual cases, they were not considered for inclusion in the dashboard as our primary objective was to offer population-level data that complement rather than supplant the conventional OFR case review model.

Apprehensions and Foreseen Challenges

Although participants were positive about the potential of the dashboard to enhance the OFR process, there were a few apprehensions. A major concern was the risk of unintentional identification of decedents in smaller counties, where there are fewer overdose deaths:

> *I've been aware of a couple different cases in relatively small communities where all the data says one thing, and of course, as a small community, we know exactly who we're talking about.* [P15]

> *I think one [concern] would be that the information might be too identifiable, especially for small communities.* [P8]

Participants discussed the ethics of displaying data that might be inaccurate or that could be misused (eg, by law enforcement) to target at-risk individuals:

> *...that it has inaccurate and bad data. And that it is used for evil rather than for good...That it's not used for bad downstream consequences kind of thing.* [P6]

Finally, participants raised the risk of misinterpreting data, noting that, while OFRs have expertise in studying individual histories of decedents to formulate recommendations, they are less familiar with analyzing population-level statistics. Some voiced reservations about OFR teams' data literacy and their ability to draw appropriate inferences from such quantitative data. For instance, 1 participant gave an example of how a decrease in emergency medical events could be erroneously interpreted as a reduction in overdoses when it might only reflect fewer 911 calls:

> *That [error] where you have a number and you think it means one thing, but it means another thing...You have measured something, but not the thing that you are taking that thing to be.* [P1]

Others commented on the potential downstream consequences of misinterpreting data, which could manifest as inappropriate or even detrimental recommendations:

> *We've seen this trend in our data. That probably means X, Y, Z. And you might be right. You might be very wrong, and the data might be used to justify a policy or programmatic intervention that could in fact exacerbate it.* [P17]

Helping users interpret data accurately was deemed by participants as a critical consideration for the dashboard. Equally important was not to inundate OFRs with even more (population-level) data that teams may lack the bandwidth or data literacy skills to act upon. These insights underscore the need to craft intuitive data visualizations that can be comprehended accurately with minimal effort. Moreover, such

XSL•FO
RenderX

displays should actively guide OFR teams into making valid inferences from the data presented.

Touchpoints Selection

Our observations point to a longstanding limitation of current OFR practices, which focus on reviewing a handful of overdose cases at every meeting. OFR experts appeared to recognize the shortcomings of this model when pitted against the sheer volume of overdoses. Simultaneously, participants expressed strong interest in accessing additional data sets that would paint a broader picture of overdose risk factors and touchpoints in their community, provided that these data were consistently coded, intuitively summarized, and presented in a manner that did not overburden review teams.

Among the data emphasized by participants, touchpoints emerged as particularly actionable as they represent system interactions preceding overdose events. For instance, the proportion of decedents who used various touchpoints offers predictive power to identify the most effective points within the system for targeting at-risk individuals with prevention services. Moreover, understanding the typical time window between a touchpoint and an overdose event, along with the frequency of touchpoint use, can assist in designing interventions, including their timing and regularity.

Drawing on the insights of the expert panel and data availability in Indiana, we incorporated 5 touchpoint types into the dashboard: jail bookings, prison releases, visits to the ED, encounters with EMS, and prescriptions for controlled substances (eg, opioid analgesics). We excluded ED and EMS encounters occurring within a 24-hour window of death as those are likely to represent interactions directly related to the overdose event as opposed to potential touchpoints for prevention purposes. Interactions with both justice and medical systems were identified as key by the expert panel. Prescriptions for scheduled drugs, such as opioid analgesics, were included as touchpoints due to their established association with overdose risk [25]. We also included the dispensation of buprenorphine prescriptions as a touchpoint in the initial dashboard design. However, concerns were raised that singling medication for opioid use disorder as a separate touchpoint could cause it to be misconstrued as a causal risk factor for overdose. Consequently, buprenorphine data were merged and included among the general prescription dispensation touchpoint for scheduled drugs. Table 1 provides a summary of these touchpoints as highlighted by participants and featured in the dashboard. Although interactions with residential treatment services were identified as an important touchpoint by participants, related data are not centrally tracked by the state and, hence, were not available for inclusion in the dashboard. Moreover, social determinants of health are not currently included despite their relevance as the dashboard was intended to prioritize opportunities for immediate as opposed to upstream prevention. Case-specific attributes were also not considered for inclusion because they would be redundant to the traditional OFR case review process.

XSL•FO
RenderX

Table 1. Data types and attributes as identified by experts and featured in the dashboard.

Data type and attribute	Identified by expert panel?	Included in dashboard?
Touchpoint		
Jail booking	No	Yes
Release from prison	Yes	Yes
Visit to the ED[a]	Yes	Yes
Encounter with EMS[b]	Yes	Yes
Interaction with residential treatment services	Yes	No
Prescription dispensation for scheduled drugs, including opioid analgesics and MOUD[c]	No	Yes
Social determinants		
Demographics	Yes	No
Educational level	Yes	No
Poverty	Yes	No
Access to harm reduction services	Yes	No
Housing	Yes	No
Access to transportation	Yes	No
Adverse childhood experiences	Yes	No
Case-specific attributes		
Toxicology report	Yes	No
Next-of-kin interviews	Yes	No
Circumstances of death (eg, body position and presence of witnesses)	Yes	No

[a]ED: emergency department.

[b]EMS: emergency medical services.

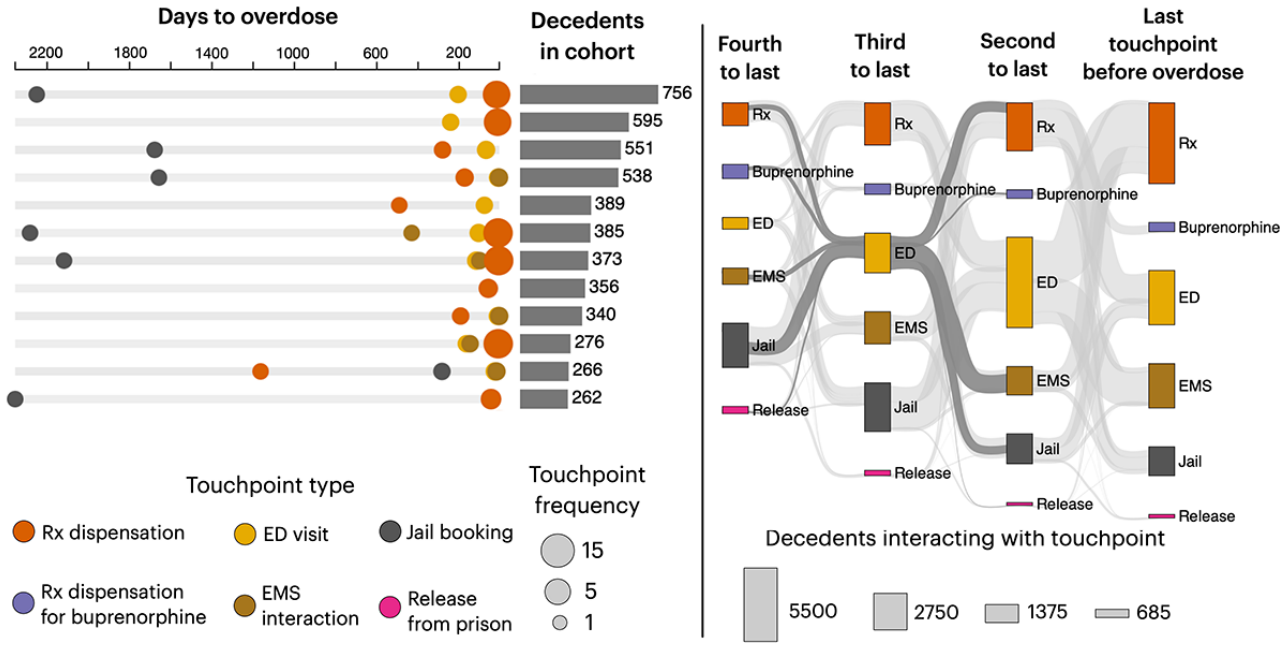
[c]MOUD: medication for opioid use disorder.

Initial Visualization Attempts

Our initial visualization focused on timelines, illustrating cohorts of decedents who exhibited similar patterns of touchpoints before overdosing. For example, in Figure 1 (left), each row represents hundreds of decedents who exhibited a similar touchpoint sequence (eg, jail booking followed by one or more ED visits and then a series of prescriptions). This particular visualization was inspired by OFR teams' use of timelines to represent the histories of individuals discussed during case reviews. However, these initial visualizations received mixed reviews from the expert panel—while they were considered appealing and "interesting," the focus on cohorts was seen as providing excessive detail for OFRs. This feedback was used to revise the visualizations and develop a final dashboard.

XSL•FO
RenderX

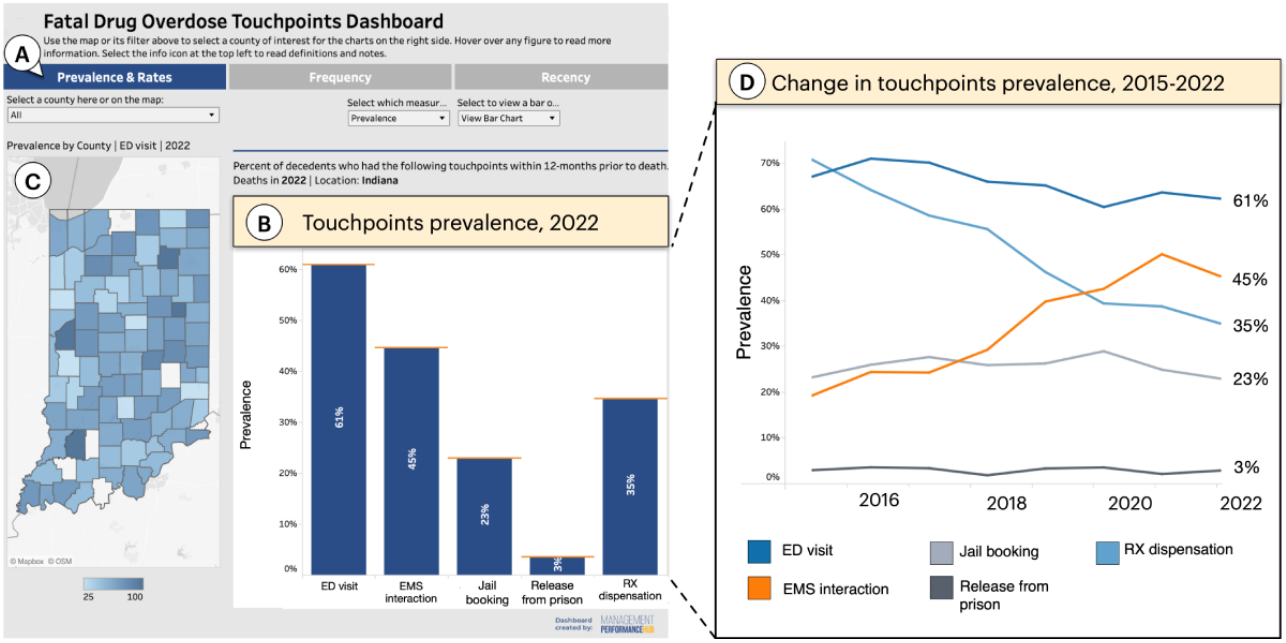
Figure 1. A total of 2 initial visual representations of touchpoints in Indiana (aggregate data from 2015 to 2022). On the left, a timeline-based visualization illustrates the cohorts of decedents with distinct sequences of touchpoints. The visualization depicts the average number of days to fatal overdose (circle position) and frequency of interaction with a touchpoint (circle diameter). For example, the first row shows 756 individuals who experienced a jail booking approximately 6 years before overdose, followed by a sequence of emergency department (ED) visits and medical prescription (Rx) dispensations, the last of which typically occurred approximately 200 and 90 days before overdose, respectively. A Sankey diagram (right) displays the temporal ordering of (up to 4) touchpoints but without showing durations. EMS: emergency medical services.



Final Dashboard

The dashboard consists of 3 primary displays (Figure 2A) showing the prevalence and rates, frequency, and recency for the 5 touchpoints. The dashboard can be accessed at the MPH website [45].

Figure 2. The final dashboard showing overall touchpoints prevalence in Indiana. (A) Buttons enable the user to switch among 4 measures: prevalence, rates, frequency, and recency of touchpoints. (B) The selected measure is visualized here as a bar chart comparing touchpoint prevalence (ie, the percentage of decedents who used each of the 5 touchpoints). (C) A map shows touchpoint prevalence (in this case for emergency department [ED] visits) by county, where darker shades of blue indicate higher prevalence. (D) As an alternative to the bar chart, a line graph allows users to observe how the prevalence of the touchpoints changes from year to year. EMS: emergency medical services; Rx: medical prescription.



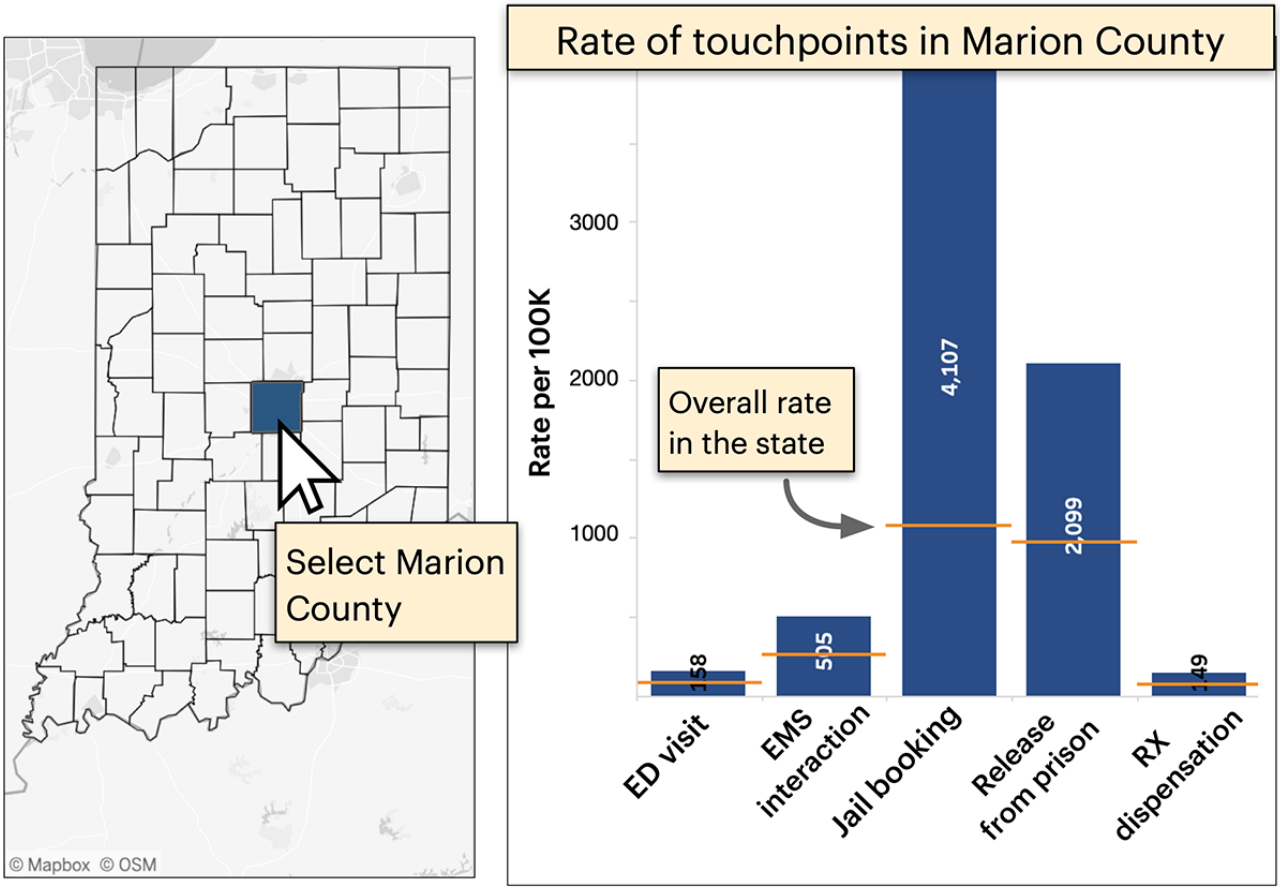
Prevalence and Rates

By default, the dashboard displays touchpoint *prevalence*, depicting the percentage of decedents who used various touchpoints in the 12 months preceding overdose. For instance, in 2022, the highest-prevalence touchpoint was the ED, with 61% of individuals who overdosed in Indiana having visited the ED within a year before dying (Figure 2B). The user can

XSL•FO
RenderX

also see the change in prevalence over time. For example, the data show that the prevalence of ED visits decreased over time, whereas the proportion of decedents who use EMS increased >2 times between 2015 and 2022 (Figure 2D). In addition to showing state levels, the dashboard can break down the data by county. For instance, the user can see the prevalence of ED visits in different counties on a map (Figure 2C). Notably, the map shows 4 counties in which practically all decedents had visited the ED a year before their overdose. The map can also be used to filter the bar or line graph displays. For example, clicking on Marion County, the most populous in Indiana, updates the display to show statistics for Marion only (Figure 3).

Figure 3. Rates showing the fraction of individuals who experienced a fatal overdose for every 100,000 people who use a touchpoint (right). A map allows the user to filter the data by county, in this example, to show rates for Marion County only. Orange dash marks depict the state average for context. ED: emergency department; EMS: emergency medical services; Rx: medical prescription.



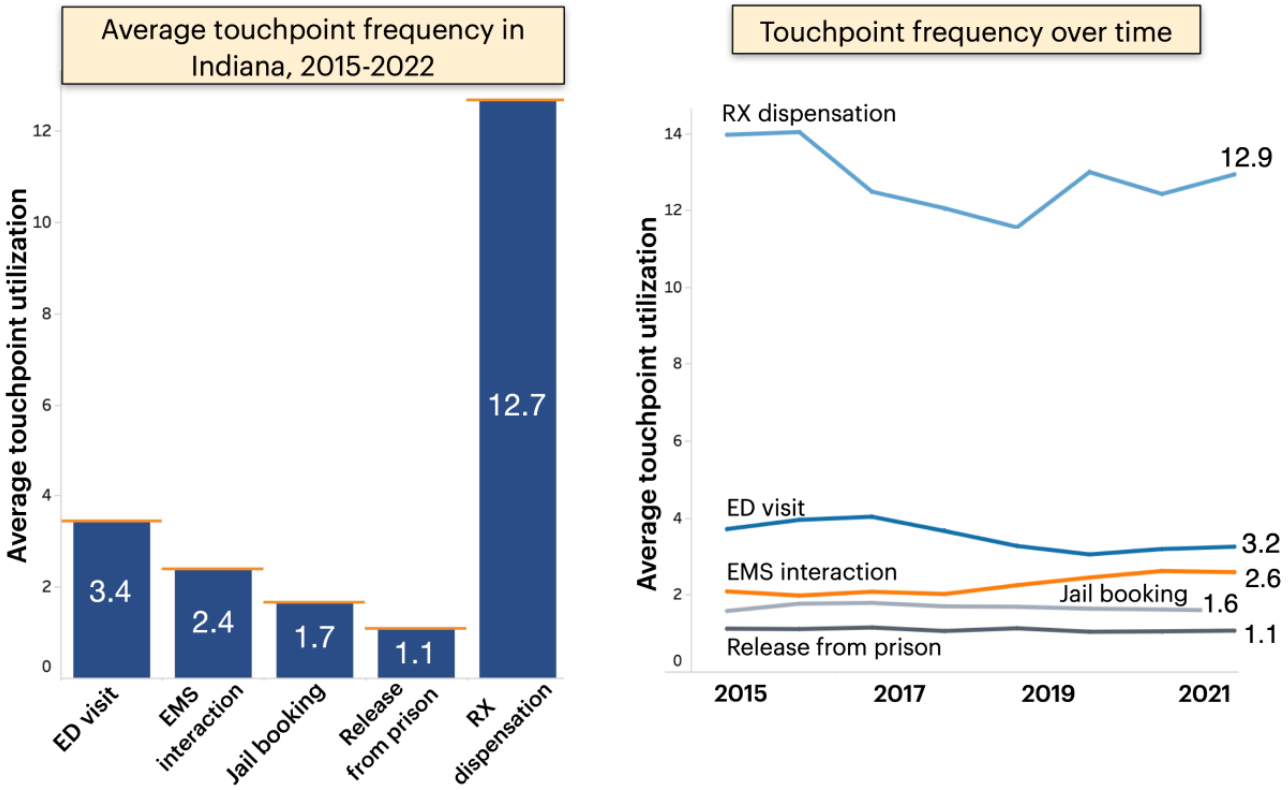
In addition to prevalence, the dashboard visualizes the *rate* of touchpoints among decedents. These rates depict the number of fatal overdose cases per 100,000 individuals who typically use services such as the ED. Unlike prevalence, which indicates the likelihood of a decedent using a touchpoint, rates reveal the probability of a fatal overdose after using 1 of the 5 legal or medical touchpoints included in the dashboard. Both measures are important for resource allocation—while prevalence helps users identify touchpoints with the broadest reach, rates can reveal more "efficient" touchpoints for targeted interventions. For example, consider jail bookings and releases from prison (Figure 3 [right]), which exhibit the highest rates among touchpoints in Marion County. This offers a high-specificity opportunity to focus on individuals at a greater risk of overdosing despite these touchpoints exhibiting relatively moderate to low prevalence at the state level (23% and 3%, respectively, as depicted in Figure 2 [left]).

Touchpoint Frequency

The second display summarizes the average number of interactions a decedent had with a touchpoint in the year preceding their overdose (Figure 4). Notably, the most frequently used touchpoint in the state is medical prescription (Rx) dispensation for controlled substances, such as an opioid analgesic (12.7 events on average at the time of writing). The user can also see how this frequency changes yearly (Figure 4 [right]). The line graph shows relatively stable use for ED, EMS, and criminal justice services, with the average number of Rx dispensations trending down slightly.

XSL•FO
RenderX

Figure 4. Average number of interactions with the 5 touchpoints from 2015 to 2022 (left) alongside a year-by-year breakdown. ED: emergency department; EMS: emergency medical services; Rx: medical prescription.

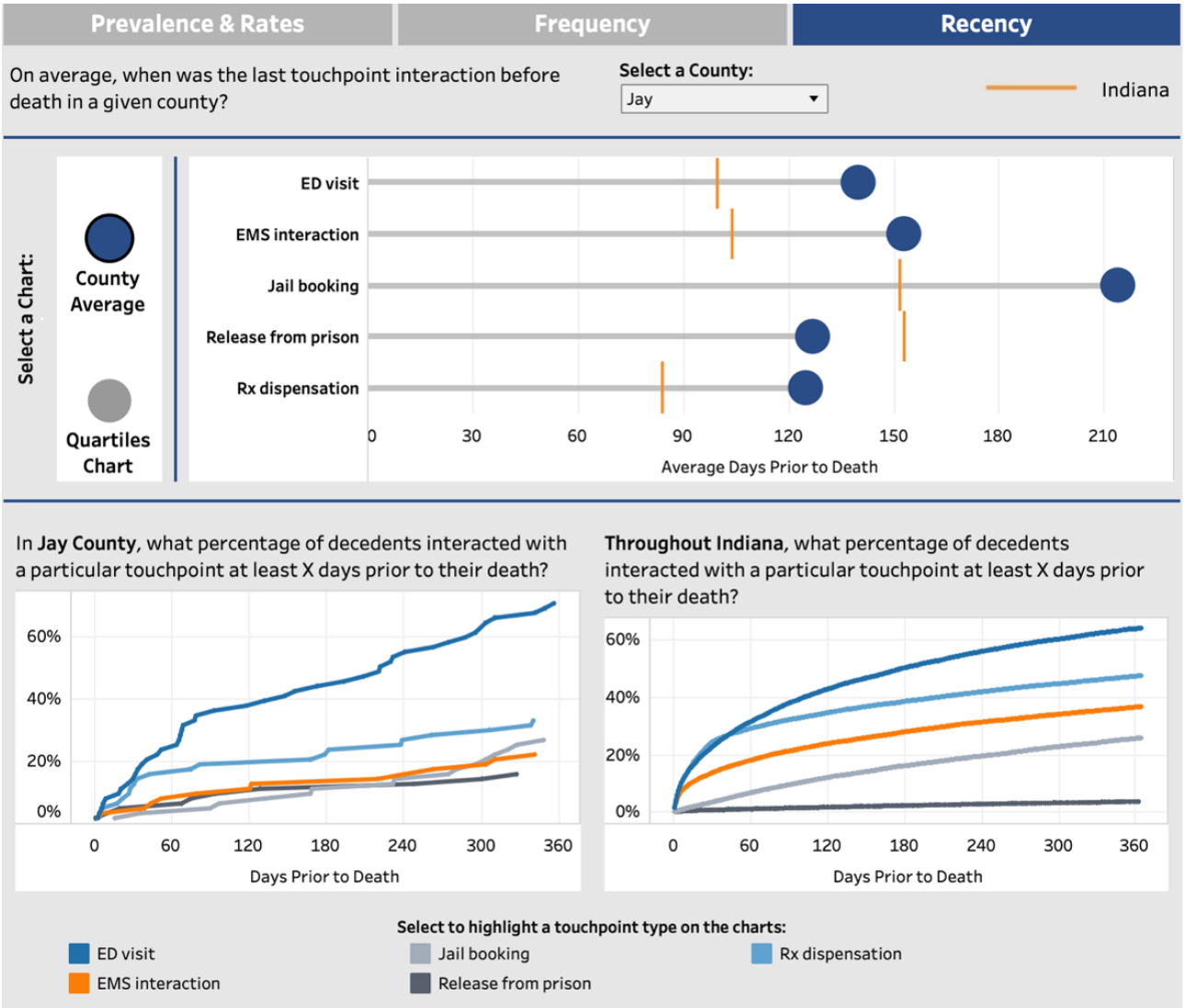


Recency

The timing of interaction with services was identified as a key factor for OFRs. Accordingly, the *recency* display illustrates the typical time intervals between final touchpoints and overdose events (Figure 5). The top features a "lollipop" chart depicting the number of days on average between the most recent interaction and the overdose (Figure 5 [top]). In this example, jail bookings in Jay County (selectable by the user) occur approximately 210 days on average before a fatal overdose compared to approximately 150 days for the entirety of Indiana. Conversely, releases from prison tend to happen approximately 120 days before the overdose, closer relative to the state average. The bottom visualizations show a curve for each touchpoint representing the cumulative percentage of individuals who could have been engaged at various time points relative to their time of death. In this case, approximately 27% of decedents in Jay County could have been engaged through an Rx dispensation touchpoint 30 days before an overdose.

XSL•FO
RenderX

Figure 5. The average time gap between the final interaction and overdose events across different touchpoints (top). The lower section comprises 2 charts demonstrating the cumulative reach of touchpoints at varying time intervals, comparing the selected county (bottom left) with the state average (bottom right). ED: emergency department; EMS: emergency medical services; Rx: medical prescription.

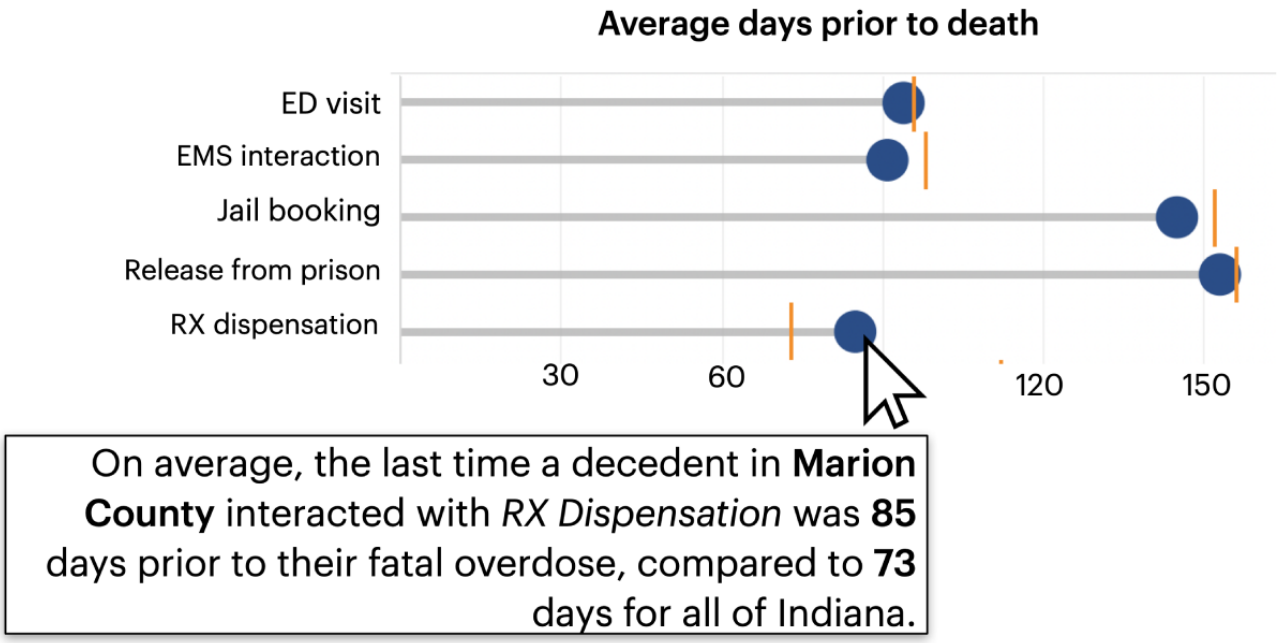


Aiding Data Interpretation

One concern that emerged during the focus groups regarded OFR teams' ability to interpret population-level statistics. To aid users in making sense of these data, the dashboard provides tooltips in the form of short text annotations that explain the interpretation of each visualization. For instance, in the recency chart, the text clarifies that the points depict the average number of days between a touchpoint and an overdose event (Figure 6).

XSL•FO
RenderX

Figure 6. Tooltips appear throughout the dashboard to promote accurate data interpretation. ED: emergency department; EMS: emergency medical services; Rx: medical prescription.



Initial Evaluation Results

We invited 3 OFR experts to review and provide feedback on the dashboard. They commented on features they thought were beneficial. They also provided suggestions on how to ensure dashboard integration into OFR practices. One of the notable strengths of the dashboard was its comprehensive data coverage, a feature that was highly appreciated by all participants. They specifically praised the breakdown of touchpoints on a county basis, a level of granularity that is often lacking in existing dashboards. The inclusion of small counties, the data on which can be especially difficult to obtain, was recognized as a significant advantage. Participants also appreciated the ability to compare different counties through the map, along with the ability to juxtapose county-specific data against state averages.

Among the various visualizations, the recency chart (referred to as the "timeline") stood out for its depiction of events leading up to overdoses. Participants thought that these temporal data, which can be difficult to obtain at the population level, can help in tailoring interventions:

It is interesting to see this [chart], and to know what can be done with data. We can check the timeline and help implement a strategy. Through these strategies, we can outline short, medium, and long-term goals.

In thinking about how the dashboard might complement existing OFR practices, participants highlighted its usefulness in guiding case selection for review and helping OFRs build a representative case profile. One participant specifically noted the potential of the dashboard in conducting "community data review" to explore "what is going on in my community." Moreover, the dashboard's availability on a publicly accessible URL was lauded as "a wonderful resource," extending its value to audiences beyond OFRs. The discussion opened the door for offering some form of training or educational support to OFR members, equipping review teams with skills to interpret quantitative data. One participant suggested the addition of a "demo video to help interpret and apply the data." Another suggested the need to specifically focus on OFR facilitators as crucial personnel for communicating data insights to review teams:

> *I don't think they [members of the review teams], will be able to fully understand the data, so training the facilitator will be key.*

Discussion

Principal Findings

OFR teams are proliferating in the United States, becoming an important public health tool to combat the drug overdose crisis. Traditional fatality reviews, often limited to a few cases, do not fully capture the broader overdose trends, especially in communities with numerous drug-related fatalities. This research aimed to enhance OFR data use by addressing data access barriers, identifying information needs, and creating actionable visualizations of population-level overdose data.

Our findings shed light on challenges that OFR teams face in accessing timely data, frequently impeded by legal constraints. When available, these data can often be inconsistent, for example, in the coding of events and classification of services. Despite these challenges, OFR teams seemed keen on incorporating a wider range of data into their review to better understand the factors contributing to overdose risks in their communities. Notably, the expert panel highlighted several key touchpoints, including incarcerations, interactions with substance treatment services, and visits to medical facilities such as EDs.

Some of these touchpoints have been previously recognized as opportunities for delivering prevention services [25,46,47]. For example, the time window following a prison release has been identified as a particularly critical and risky period, making this touchpoint a highly specific and valuable opportunity for

XSL•FO
RenderX

administering prevention services [48-50]. However, effectively sharing these data insights with OFRs remains a challenge. Our findings suggest that a dashboard linking state administrative and mortality data could effectively provide local OFRs with insights on the timing and distribution of touchpoints. To explore this potential, we partnered with the Indiana state government and developed a dashboard that collates and visualizes data on 5 touchpoints at the county level, enabling OFR teams to see statistics and patterns on events that *precede* fatal overdoses in their community. To our knowledge, this is the first system to automatically analyze touchpoint characteristics and offer (near) real-time visualizations of their prevalence, frequency, and timing tailored to the local scale of OFR teams. In designing the dashboard, we specifically focused on this user group and prioritized actionable data that shed light on local prevention opportunities. The developed touchpoint dashboard stands in contrast to earlier dashboards for opioid prescription and overdose data, which are meant for the public or nonspecified stakeholders.

Our OFR expert panel suggested that one of the most crucial pieces of information is the timing of touchpoints—specifically, the average duration between an individual's last encounter and their overdose. The dashboard prominently features these data in a lollipop chart comparing the *recency* of various touchpoints. In addition, we incorporated displays of touchpoint prevalence and rates, providing insights into the reach of touchpoints and the specificity they afford for targeting individuals who are at high risk of overdose. The dashboard purposely uses familiar visualizations, including bar and line graphs and choropleth maps, to appeal to review teams who may be novice visualization users [51]. Importantly, the dashboard breaks down these statistics at the county level, aligning with how OFRs are organized in Indiana. By visualizing data "close to home," we aimed to improve the actionability of the dashboard [52]. However, users can easily compare county data to state averages or those of other similar counties.

Our initial evaluations show promise for the dashboard's usefulness. However, successfully integrating the dashboard into OFR practices will likely require training for OFR members, many of whom lack expertise in data analysis—a point that was notably underscored by the expert panel. In particular, teams may need educational support in how to interpret population-level features, such as the difference between the prevalence and rates of touchpoints. Regular meetings with OFR users could also help uncover usability issues and gauge dashboard adoption by review teams.

While the dashboard offers detailed insights into community touchpoints, it omits data on social determinants such as race, educational level, and access to housing and harm reduction services. These factors can be important for understanding overdose risks, as per our expert panel and research findings [53,54]. Future versions of the dashboard could incorporate local statistics on these risk factors. Furthermore, it is possible to expand the current list of touchpoints to include specific events associated with social determinants, such as loss of housing or employment. These additional touchpoints could offer further intervention avenues to disrupt pathways from marginalization to overdose [55]. Another limitation is that, while the dashboard includes critical touchpoints such as ED and EMS encounters, these events currently lack classification. Adding a breakdown of these touchpoints, for example, by distinguishing between substance-related versus other EMS encounters, could enable OFR teams to further tailor their recommendations.

The experts interviewed also sought demographic breakdowns of touchpoint data, in part to ensure that diverse populations would benefit from interventions at touchpoints. Unfortunately, this feature was not included in the current dashboard due to reidentification risks, particularly in rural areas that have fewer overdoses. In the future, the dashboard could be modified to provide a demographic breakdown of touchpoints at the aggregate (eg, state) level to substantially decrease the risk of reidentification instead of withholding these data altogether. To further protect individual confidentiality, which was a key concern of our expert panel, the dashboard presents data as percentages (eg, the proportion of decedents who were released from prison within a year before their overdose) and rates. Withholding the actual counts for events helps prevent the inference of individual identity in places where those counts are low. The dashboard provides a visual warning for statistics based on <20 cases, cautioning users against drawing strong conclusions from small samples. Future work could use more advanced privacy-preserving techniques [56,57], thus allowing for the display of a wider range of attributes without jeopardizing anonymity.

Although our dashboard is specific to Indiana, we believe that the approach could be adapted for other US states and localities. This expansion requires access to overdose mortality records that can be algorithmically cross-referenced with other administrative data sets. Many states already have data infrastructure for such linked analyses [58,59]. We estimate that the development and maintenance of the dashboard over 2 years will require approximately 350 personnel hours assuming the availability of data. The prevalence of overdose dashboards [39,60] indicates both the technical feasibility of creating such tools and the interest in them from the public health community. Our research demonstrates that dashboards can go beyond surveillance to directly visualize actionable prevention opportunities.

Conclusions

OFRs can play a crucial public health role in understanding overdose cases and recommending prevention strategies. This study explored the potential for enhancing these reviews with population-level data for broader, quantitative insights into risk factors. Following a user-centered design process, we developed a dashboard that tracks and visualizes decedents' encounters with medical and justice systems at the county level. Although initially designed for Indiana, the dashboard can be adapted to other localities, leveraging administrative and mortality data typically collected by local governments. Preliminary evaluation shows the potential utility of the dashboard for analysis and case selection but emphasizes the need for training OFR members in data interpretation and decision-making.

XSL•FO
RenderX

Acknowledgments

The authors thank the Indiana Department of Health; the Department of Homeland Security; the Indiana State Police; the Professional Licensing Agency; the Department of Correction; and the Governor's Drug Prevention, Treatment, and Enforcement Office for providing data for use in the dashboard and for feedback. This research was supported by the Helping to End Addiction Long-term Initiative through the Data 2 Action program (award R61DA057660). The work was also supported in part by Research Jam, the Indiana Clinical and Translational Sciences Institute Patient Engagement Core through an award from the National Institutes of Health, National Center for Advancing Translational Sciences Clinical and Translational Science Award (award UM1TR004402). The authors acknowledge Michelle Tran, BS, for her contributions to this work. The contents of this publication are solely the responsibility of the authors and do not necessarily represent the official views of the National Institutes of Health or participating organizations.

Data Availability

The video recording data sets generated during and analyzed during this study are not publicly available to protect participants' confidentiality but are available from the corresponding author on reasonable request. The aggregate touchpoints data sets are available in the Management Performance Hub (MPH) data repository [61].

Authors' Contributions

AS and LAG analyzed the data and wrote the manuscript. BPC conducted the focus group and analyzed the data. SEW directed the human-centered research process and provided manuscript revisions. KC, JC, and TB developed the dashboard and the touchpoints data integration process. KS, ALD, BR, MCA, and KR edited and provided manuscript revisions. BR, MCA, and KR conceptualized the study and designed the research.

Conflicts of Interest

None declared.

References

1. Brady JE, Giglio R, Keyes KM, DiMaggio C, Li G. Risk markers for fatal and non-fatal prescription drug overdose: a meta-analysis. Inj Epidemiol 2017 Dec 7;4(1):24 [FREE Full text] [doi: 10.1186/s40621-017-0118-7] [Medline: 28762157]
2. Davoli M, Perucci CA, Forastiere F, Doyle P, Rapiti E, Zaccarelli M, et al. Risk factors for overdose mortality: a case-control study within a cohort of intravenous drug users. Int J Epidemiol 1993 Apr;22(2):273-277. [doi: 10.1093/ije/22.2.273] [Medline: 8505184]
3. Webster LR. Risk factors for opioid-use disorder and overdose. Anesth Analg 2017 Nov;125(5):1741-1748. [doi: 10.1213/ANE.0000000000002496] [Medline: 29049118]
4. Lyons RM, Yule AM, Schiff D, Bagley SM, Wilens TE. Risk factors for drug overdose in young people: a systematic review of the literature. J Child Adolesc Psychopharmacol 2019 Aug 01;29(7):487-497 [FREE Full text] [doi: 10.1089/cap.2019.0013] [Medline: 31246496]
5. Yamamoto A, Needleman J, Gelberg L, Kominski G, Shoptaw S, Tsugawa Y. Association between homelessness and opioid overdose and opioid-related hospital admissions/emergency department visits. Soc Sci Med 2019 Dec;242:112585 [FREE Full text] [doi: 10.1016/j.socscimed.2019.112585] [Medline: 31634808]
6. Binswanger IA, Blatchford PJ, Mueller SR, Stern MF. Mortality after prison release: opioid overdose and other causes of death, risk factors, and time trends from 1999 to 2009. Ann Intern Med 2013 Nov 05;159(9):592-600 [FREE Full text] [doi: 10.7326/0003-4819-159-9-201311050-00005] [Medline: 24189594]
7. Bukten A, Stavseth MR, Skurtveit S, Tverdal A, Strang J, Clausen T. High risk of overdose death following release from prison: variations in mortality during a 15-year observation period. Addiction 2017 Aug 07;112(8):1432-1439. [doi: 10.1111/add.13803] [Medline: 28319291]
8. Barboza G, Angulski K, Hines L, Brown P. Variability in opioid-related drug overdoses, social distancing, and area-level deprivation during the COVID-19 pandemic: a Bayesian spatiotemporal analysis. J Urban Health 2022 Oct 06;99(5):873-886 [FREE Full text] [doi: 10.1007/s11524-022-00675-x] [Medline: 36068454]
9. Forati AM, Ghose R, Mantsch JR. Examining opioid overdose deaths across communities defined by racial composition: a multiscale geographically weighted regression approach. J Urban Health 2021 Aug 06;98(4):551-562 [FREE Full text] [doi: 10.1007/s11524-021-00554-x] [Medline: 34231120]
10. Ghose R, Forati AM, Mantsch JR. Impact of the COVID-19 pandemic on opioid overdose deaths: a spatiotemporal analysis. J Urban Health 2022 Apr 18;99(2):316-327 [FREE Full text] [doi: 10.1007/s11524-022-00610-0] [Medline: 35181834]

XSL•FO
RenderX

11. Brady JE, DiMaggio CJ, Keyes KM, Doyle JJ, Richardson LD, Li G. Emergency department utilization and subsequent prescription drug overdose death. Ann Epidemiol 2015 Aug;25(8):613-9.e2 [FREE Full text] [doi: 10.1016/j.annepidem.2015.03.018] [Medline: 25935710]
12. Choe K, Zinn E, Lu K, Hoang D, Yang LH. Impact of COVID-19 pandemic on chronic pain and opioid use in marginalized populations: a scoping review. Front Public Health 2023 Apr 17;11:1046683 [FREE Full text] [doi: 10.3389/fpubh.2023.1046683] [Medline: 37139395]
13. Foglia R, Kline A, Cooperman NA. New and emerging opioid overdose risk factors. Curr Addict Rep 2021 Apr 22;8(2):319-329 [FREE Full text] [doi: 10.1007/s40429-021-00368-6] [Medline: 33907663]
14. Durfee M, Parra JM, Alexander R. Child fatality review teams. Pediatr Clin North Am 2009 Apr;56(2):379-387. [doi: 10.1016/j.pcl.2009.01.004] [Medline: 19358922]
15. Rebbert-Franklin K, Haas E, Singal P, Cherico-Hsii S, Baier M, Collins K, et al. Development of Maryland local overdose fatality review teams: a localized, interdisciplinary approach to combat the growing problem of drug overdose deaths. Health Promot Pract 2016 Jul 18;17(4):596-600. [doi: 10.1177/1524839916632549] [Medline: 27091609]
16. Jones C, Bracewell K, Clegg A, Stanley N, Chantler K. Domestic homicide review committees' recommendations and impacts: a systematic review. Homicide Stud 2022 Apr 08;28(1):78-98 [FREE Full text] [doi: 10.1177/10887679221081788]
17. Parrish JW, Schnitzer PG, Lanier P, Shanahan ME, Daniels JL, Marshall SW. Classification of maltreatment-related mortality by child death review teams: how reliable are they? Child Abuse Negl 2017 May;67:362-370. [doi: 10.1016/j.chiabu.2017.03.003] [Medline: 28365427]
18. Evaluation profile for overdose fatality reviews. Centers for Disease Control and Prevention. URL: https://www.cdc.gov/overdose-prevention/media/pdfs/OD2A_EvalProfile_OverdoseFatalityReviews_508.pdf [accessed 2024-01-30]
19. Haas E, Truong C, Bartolomei-Hill L, Baier M, Bazron B, Rebbert-Franklin K. Local overdose fatality review team recommendations for overdose death prevention. Health Promot Pract 2019 Jul 11;20(4):553-564. [doi: 10.1177/1524839918797617] [Medline: 30205737]
20. Keen C, Young JT, Borschmann R, Kinner SA. Non-fatal drug overdose after release from prison: a prospective data linkage study. Drug Alcohol Depend 2020 Jan 01;206:107707. [doi: 10.1016/j.drugalcdep.2019.107707] [Medline: 31757517]
21. McLeod KE, Slaunwhite AK, Zhao B, Moe J, Purssell R, Gan W, et al. Comparing mortality and healthcare utilization in the year following a paramedic-attended non-fatal overdose among people who were and were not transported to hospital: a prospective cohort study using linked administrative health data. Drug Alcohol Depend 2021 Jan 01;218:108381. [doi: 10.1016/j.drugalcdep.2020.108381] [Medline: 33158663]
22. Shefner RT, Sloan JS, Sandler KR, Anderson ED. Missed opportunities: arrest and court touchpoints for individuals who fatally overdosed in Philadelphia in 2016. Int J Drug Policy 2020 Apr;78:102724. [doi: 10.1016/j.drugpo.2020.102724] [Medline: 32279054]
23. Stoové MA, Dietze PM, Jolley D. Overdose deaths following previous non-fatal heroin overdose: record linkage of ambulance attendance and death registry data. Drug Alcohol Rev 2009 Jul 06;28(4):347-352. [doi: 10.1111/j.1465-3362.2009.00057.x] [Medline: 19594787]
24. Tseregounis IE, Henry SG. Assessing opioid overdose risk: a review of clinical prediction models utilizing patient-level data. Transl Res 2021 Aug;234:74-87 [FREE Full text] [doi: 10.1016/j.trsl.2021.03.012] [Medline: 33762186]
25. Larochelle MR, Bernstein R, Bernson D, Land T, Stopka TJ, Rose AJ, et al. Touchpoints - opportunities to predict and prevent opioid overdose: a cohort study. Drug Alcohol Depend 2019 Nov 01;204:107537 [FREE Full text] [doi: 10.1016/j.drugalcdep.2019.06.039] [Medline: 31521956]
26. Ray B, Christian K, Bailey T, Alton M, Proctor A, Haggerty J, et al. Antecedents of fatal overdose in an adult cohort identified through administrative record linkage in Indiana, 2015-2022. Drug Alcohol Depend 2023 Jun 01;247:109891 [FREE Full text] [doi: 10.1016/j.drugalcdep.2023.109891] [Medline: 37120921]
27. Krawczyk N, Schneider KE, Eisenberg MD, Richards TM, Ferris L, Mojtabai R, et al. Opioid overdose death following criminal justice involvement: linking statewide corrections and hospital databases to detect individuals at highest risk. Drug Alcohol Depend 2020 Aug 01;213:107997. [doi: 10.1016/j.drugalcdep.2020.107997] [Medline: 32534407]
28. Dasgupta N, Kapadia F. The future of the public health data dashboard. Am J Public Health 2022 Jun;112(6):886-888. [doi: 10.2105/ajph.2022.306871]
29. Thorpe LE, Gourevitch MN. Data dashboards for advancing health and equity: proving their promise? Am J Public Health 2022 Jun;112(6):889-892. [doi: 10.2105/ajph.2022.306847]
30. Bach B, Freeman E, Abdul-Rahman A, Turkay C, Khan S, Fan Y, et al. Dashboard design patterns. IEEE Trans Vis Comput Graph 2023;29(1):342-352. [doi: 10.1109/tvcg.2022.3209448]
31. Few S. Information Dashboard Design: The Effective Visual Communication of Data. Sebastopol, CA: O'Reilly Media; 2006.
32. Gleeson J, Kitchin R, McCarthy E. Dashboards and public health: the development, impacts, and lessons from the Irish government COVID-19 dashboards. Am J Public Health 2022 Jun;112(6):896-899. [doi: 10.2105/ajph.2022.306848]
33. Lechner B, Fruhling A. Towards public health dashboard design guidelines. In: Proceedings of the 1st International Conference, HCIB 2014, Held as Part of HCI International on HCI in Business. 2014 Presented at: HCIB '14; June 22-27,

XSL•FO
RenderX

2014; Crete, Greece p. 49-59 URL: https://link.springer.com/chapter/10.1007/978-3-319-07293-7_5 [doi: 10.1007/978-3-319-07293-7_5]
34. Amarasinghe I, Hernandez-Leo D, Michos K, Vujovic M. An actionable orchestration dashboard to enhance collaboration in the classroom. IEEE Trans Learn Technol 2020 Oct 1;13(4):662-675. [doi: 10.1109/tlt.2020.3028597]
35. Lock O, Bednarz T, Leao SZ, Pettit C. A review and reframing of participatory urban dashboards. City Cult Soc 2020 Mar;20:100294. [doi: 10.1016/j.ccs.2019.100294]
36. Setlur V, Correll M, Satyanarayan A, Tory M. Heuristics for supporting cooperative dashboard design. IEEE Trans Vis Comput Graph 2024 Jan;30(1):370-380. [doi: 10.1109/TVCG.2023.3327158] [Medline: 37883257]
37. Anderson JA, Demeter N, Pasquires M, Wirtz S. Using the CA opioid overdose surveillance dashboard to track opioid overdose deaths. Online J Public Health Inform 2019 May 30;11(1):e449 [FREE Full text] [doi: 10.5210/ojphi.v11i1.9938]
38. Safranek CW, Feitzinger L, Joyner AK, Woo N, Smith V, Souza ED, et al. Visualizing opioid-use variation in a pediatric perioperative dashboard. Appl Clin Inform 2022 Mar 23;13(2):370-379 [FREE Full text] [doi: 10.1055/s-0042-1744387] [Medline: 35322398]
39. Wu E, Villani J, Davis A, Fareed N, Harris DR, Huerta TR, et al. Community dashboards to support data-informed decision-making in the HEALing communities study. Drug Alcohol Depend 2020 Dec 01;217:108331 [FREE Full text] [doi: 10.1016/j.drugalcdep.2020.108331] [Medline: 33070058]
40. Abras C, Maloney-Krichmar D, Preece J. User-centered design. In: Ghaoui C, editor. Encyclopedia of Human Computer Interaction. Thousand Oaks, CA: Sage Publications; 2004:445-456.
41. Vredenburg K, Mao JY, Smith PW, Carey T. A survey of user-centered design practice. In: Proceedings of the 2002 SIGCHI Conference on Human Factors in Computing Systems. 2002 Presented at: CHI '02; April 20-25, 2002; Minneapolis, MN p. 471-478 URL: https://dl.acm.org/doi/10.1145/503376.503460
42. Muller MJ, Haslwanter JH, Dayton T. Participatory practices in the software lifecycle. In: Helander MG, Landauer TK, Prabhu PV, editors. Handbook of Human-Computer Interaction. Amsterdam, The Netherlands: North Holland; 1997:255-297.
43. Schuler D, Namioka A. Participatory Design: Principles and Practices. Boca Raton, FL: CRC Press; 1993.
44. Guest G, MacQueen KM, Namey EE. Applied Thematic Analysis. Thousand Oaks, CA: Sage Publications; 2011.
45. Fatal drug overdose touchpoints. Indiana Management Performance Hub. URL: https://www.in.gov/mph/projects/fortress/ [accessed 2024-01-30]
46. Bennett AS, Elliott L. Naloxone's role in the national opioid crisis-past struggles, current efforts, and future opportunities. Transl Res 2021 Aug;234:43-57 [FREE Full text] [doi: 10.1016/j.trsl.2021.03.001] [Medline: 33684591]
47. Snyder H, Kalmin MM, Moulin A, Campbell A, Goodman-Meza D, Padwa H, et al. Rapid adoption of low-threshold buprenorphine treatment at California emergency departments participating in the CA bridge program. Ann Emerg Med 2021 Dec;78(6):759-772 [FREE Full text] [doi: 10.1016/j.annemergmed.2021.05.024] [Medline: 34353655]
48. Binswanger IA, Nowels C, Corsi KF, Glanz J, Long J, Booth RE, et al. Return to drug use and overdose after release from prison: a qualitative study of risk and protective factors. Addict Sci Clin Pract 2012 Mar 15;7(1):3 [FREE Full text] [doi: 10.1186/1940-0640-7-3] [Medline: 22966409]
49. Cooper JA, Onyeka I, Cardwell C, Paterson E, Kirk R, O'Reilly D, et al. Record linkage studies of drug-related deaths among adults who were released from prison to the community: a scoping review. BMC Public Health 2023 May 05;23(1):826 [FREE Full text] [doi: 10.1186/s12889-023-15673-0] [Medline: 37147595]
50. Merrall EL, Kariminia A, Binswanger IA, Hobbs MS, Farrell M, Marsden J, et al. Meta-analysis of drug-related deaths soon after release from prison. Addiction 2010 Sep 05;105(9):1545-1554 [FREE Full text] [doi: 10.1111/j.1360-0443.2010.02990.x] [Medline: 20579009]
51. Grammel L, Tory M, Storey MA. How information visualization novices construct visualizations. IEEE Trans Vis Comput Graph 2010 Nov;16(6):943-952. [doi: 10.1109/TVCG.2010.164] [Medline: 20975131]
52. Ivanković D, Barbazza E, Bos V, Brito Fernandes Ó, Jamieson Gilmore K, Jansen T, et al. Features constituting actionable COVID-19 dashboards: descriptive assessment and expert appraisal of 158 public web-based COVID-19 dashboards. J Med Internet Res 2021 Feb 24;23(2):e25682 [FREE Full text] [doi: 10.2196/25682] [Medline: 33577467]
53. Kariisa M, Davis NL, Kumar S, Seth P, Mattson CL, Chowdhury F, et al. Vital signs: drug overdose deaths, by selected sociodemographic and social determinants of health characteristics - 25 states and the district of Columbia, 2019-2020. MMWR Morb Mortal Wkly Rep 2022 Jul 22;71(29):940-947 [FREE Full text] [doi: 10.15585/mmwr.mm7129e2] [Medline: 35862289]
54. Park JN, Rouhani S, Beletsky L, Vincent L, Saloner B, Sherman SG. Situating the continuum of overdose risk in the social determinants of health: a new conceptual framework. Milbank Q 2020 Sep 18;98(3):700-746 [FREE Full text] [doi: 10.1111/1468-0009.12470] [Medline: 32808709]
55. van Draanen J, Jamula R, Karamouzian M, Mitra S, Richardson L. Pathways connecting socioeconomic marginalization and overdose: a qualitative narrative synthesis. Int J Drug Policy 2023 Mar;113:103971. [doi: 10.1016/j.drugpo.2023.103971] [Medline: 36822011]
56. Agrawal R, Srikant R. Privacy-preserving data mining. In: Proceedings of the 2000 ACM SIGMOD International Conference on Management of Data. 2000 Presented at: SIGMOD '00; May 15-18, 2000; Dallas, TX p. 439-450 URL: https://dl.acm.org/doi/10.1145/335191.335438 [doi: 10.1145/342009.335438]

XSL•FO
RenderX

57. Bhattacharjee K, Chen M, Dasgupta A. Privacy-preserving data visualization: reflections on the state of the art and research opportunities. Comput Graph Forum 2020 Jul 18;39(3):675-692. [doi: 10.1111/cgf.14032]
58. Evans EA, Delorme E, Cyr K, Goldstein DM. A qualitative study of big data and the opioid epidemic: recommendations for data governance. BMC Med Ethics 2020 Oct 21;21(1):101 [FREE Full text] [doi: 10.1186/s12910-020-00544-9] [Medline: 33087123]
59. Evans EA, Delorme E, Cyr KD, Geissler KH. The Massachusetts public health data warehouse and the opioid epidemic: a qualitative study of perceived strengths and limitations for advancing research. Prev Med Rep 2022 Aug;28:101847 [FREE Full text] [doi: 10.1016/j.pmedr.2022.101847] [Medline: 35669857]
60. Waye KM, Yedinak JL, Koziol J, Marshall BD. Action-focused, plain language communication for overdose prevention: a qualitative analysis of Rhode Island's overdose surveillance and information dashboard. Int J Drug Policy 2018 Dec;62:86-93. [doi: 10.1016/j.drugpo.2018.08.010] [Medline: 30384027]
61. Fatal drug overdose touchpoints. Management Performance Hub, Indiana MPH. URL: https://www.in.gov/mph/projects/fortress/ [accessed 2024-04-29]

Abbreviations

ED: emergency department
EMS: emergency medical services
MPH: Management Performance Hub
OFR: Overdose Fatality Review
Rx: medical prescription

Edited by A Kushniruk; submitted 08.02.24; peer-reviewed by K Wheeler; comments to author 08.03.24; revised version received 24.03.24; accepted 13.05.24; published 11.06.24.

Please cite as:
Salvi A, Gillenwater LA, Cockrum BP, Wiehe SE, Christian K, Cayton J, Bailey T, Schwartz K, Dir AL, Ray B, Aalsma MC, Reda K
Development of a Real-Time Dashboard for Overdose Touchpoints: User-Centered Design Approach
JMIR Hum Factors 2024;11:e57239
URL: https://humanfactors.jmir.org/2024/1/e57239
doi:10.2196/57239
PMID:38861717

©Amey Salvi, Logan A Gillenwater, Brandon P Cockrum, Sarah E Wiehe, Kaitlyn Christian, John Cayton, Timothy Bailey, Katherine Schwartz, Allyson L Dir, Bradley Ray, Matthew C Aalsma, Khairi Reda. Originally published in JMIR Human Factors (https://humanfactors.jmir.org), 11.06.2024. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on https://humanfactors.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Designing a Smartphone-Based Pulse Oximeter for Children in South Africa (Phefumla Project): Qualitative Analysis of Human-Centered Design Workshops With Health Care Workers

Elif I Ilhan[1], MA; Lucia N Jola[1], MA; Marieke M van der Zalm[1], PhD; Mike Bernstein[2], BSc; Pierre Goussard[3], PhD; Andrew Redfern[3], MBCHB, MPhil; Anneke C Hesseling[1], PhD; Graeme Hoddinott[1,4], PhD; Eric D McCollum[1,5], MPH, MD; Carina King[6], PhD

1

2

3

4

5

6

Corresponding Author:
Carina King, PhD

Abstract

Background: Pulse oximeters noninvasively measure blood oxygen levels, but these devices have rarely been designed for low-resource settings and are inconsistently available at outpatient clinics.

Objective: The *Phefumla* project aims to develop and validate a pediatric smartphone-based pulse oximeter designed specifically for this context. We present the process of human-centered oximeter design with health care workers in South Africa.

Methods: We purposively sampled 19 health care workers from 5 clinics in Khayelitsha, Cape Town. Using a human-centered design approach, we conducted participatory workshops with four activities with health care workers: (1) they received 3D-printed prototypes of potential oximeter designs to provide feedback; (2) we demonstrated on dolls how they would use the novel oximeter; (3) they used pile sorting to rank design features and suggest additional features they desired; and (4) they designed their preferred user interface using a whiteboard, marker, and magnetized features that could be repositioned. We audio recorded the workshops, photographed outputs, and took detailed field notes. Analysis involved iterative review of these data to describe preferences, identify key design updates, and provide modifications.

Results: Participants expressed a positive sentiment toward the idea of a smartphone pulse oximeter and suggested that a pediatric device would address an important gap in outpatient care. Specifically, participants expressed a preference for the prototype that they felt enabled more diversity in the way it could be used. There was a strong tendency to prioritize pragmatic design features, such as robustness, which was largely dictated by health care worker context. They also added features that would allow the oximeter device to serve other clinical functions in addition to oxygen saturation measurement, such as temperature and respiratory rate measurements.

Conclusions: Our end user–centered rapid participatory approach led to tangible design changes and prompted design discussions that the team had not previously considered. Overall, health care workers prioritized pragmatism for pediatric pulse oximeter device design.

(JMIR Hum Factors 2024;11:e54983) doi:10.2196/54983

KEYWORDS

pediatrics; human-centered design; participatory design; pulse oximeter; South Africa; smartphone; mobile phone

Introduction

Hypoxemia, defined as an abnormally low peripheral arterial oxyhemoglobin saturation (SpO_2) of <90%, is an important risk factor for death among children with lower respiratory infections in low- and middle-income countries (LMICs) [1-4]. An estimated 7 million children were hospitalized with hypoxemic pneumonia in 2019, and in sub-Saharan Africa, 28% (95% CI 25%-35%) of children with acute respiratory diseases were hypoxemic [5]. During outpatient care, the burden of hypoxemia may be considerable, with 2019 estimates suggesting a 23.1% prevalence among children with respiratory illnesses [5]. Pulse

XSL•FO
RenderX

oximeters are medical devices that noninvasively measure SpO_2 and can therefore detect hypoxemia.

Frequent pulse oximeter use is associated with positive health outcomes such as reducing mortality rates and has been found to be cost-effective in low-resource settings [1,6]. Although oximeters are commonly used in pediatric clinical care in high-income countries, they are not consistently available in LMICs [4], especially during outpatient care where most children first access care and their illness may be more treatment responsive. The COVID-19 pandemic led to large investments being made into oxygen ecosystems, including pulse oximetry [7]; however, it did not focus on overcoming key implementation challenges for children. Pediatric pulse oximetry implementation in LMICs is restricted by barriers such as cost, lack of appropriately designed pediatric devices and probes, disruptive movements of small children, unavailability of devices, lack of training and supervision, lack of maintenance, lack of electricity, and health care provider misconceptions [8-14]. A pediatric-specialized, low-cost, smartphone-based pulse oximeter device could potentially address many of these implementation barriers and serve as a valuable tool in outpatient LMIC settings.

The *Phefumla* project aims to cocreate a low-cost, smartphone-based, reflectance pulse oximeter device for children that is optimized for LMIC outpatient contexts. Reflectance oximetry, unlike transmittance oximetry, measures the relative ratio of unabsorbed red and infrared light that is reflected off of tissues rather than through tissues to produce an estimate of SpO_2 [15]. A key source of inequities in health is access to diagnostic services, with almost half of the global population having little or no access to diagnostics [16]. Part of this inequality stems from devices designed for high-income and inpatient settings that are cost-prohibitive to purchase, sustain consumables, and maintain. To address this challenge, there is a need for a holistic framework to guide the design of medical devices so that they may be contextually appropriate for the settings in which they will be used [17].

We used a human-centered design (HCD) approach, with the aim of achieving a contextually appropriate device that meets the specific health care needs of the population [18]. The HCD method is one of many approaches to co-design and was chosen for this study given its successful application in previous global health intervention and medical device development projects [19-25]. In this paper, we describe the participatory HCD processes with health care workers (HCWs) and how this led to design changes, as an example of a rapid approach to medical device development that centers inclusion.

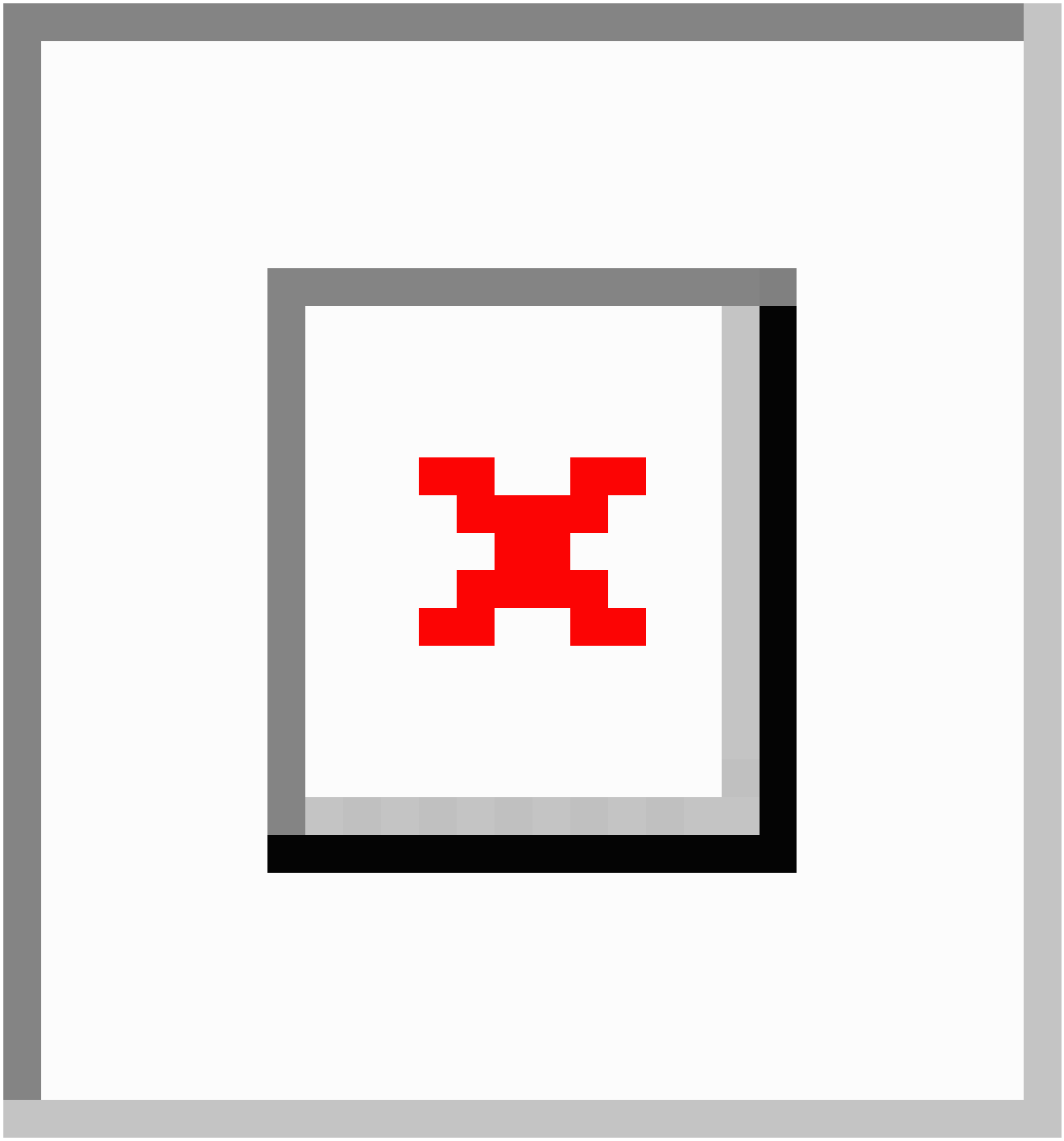
Methods

Overview

We conducted a qualitative observation study of participatory workshops that drew on the HCD approach, with HCWs in Khayelitsha, Cape Town, South Africa, from September 1-16, 2022. For these workshops, we had 3D-printed 3 prototype reflectance devices, all based on the same smartphone model being housed inside a case that would contain the oximeter sensor and additional hardware for processing (Figure 1). These prototypes were developed by the *Phefumla* team to prompt HCW reflections on the size and positioning of the sensor while keeping all other factors consistent.

XSL•FO
RenderX

Figure 1. 3D-printed *Phefumla* reflectance oximeter prototypes.



Setting

The East subdistrict of Khayelitsha is a low-income and low-resource area in Cape Town, South Africa, often referred to as a township. It has an estimated population of 450,000 people [26], who are predominantly Black African (99%) and the majority of whom live in informal housing [27]. First-tier primary health care (PHC) in South Africa is provided primarily through nurse-led clinics and community health centers, which are available within 5 km of 90% of the population and is often the first point of contact [28]. These facilities are free of cost and provide comprehensive basic services such as maternal, child, and reproductive health; HIV and tuberculosis testing and treatment; and care for noncommunicable diseases and common ailments [28]. Secondary care is delivered at district hospitals, which conduct minor procedures, and the third tier consists of tertiary hospitals that have the infrastructure, specialists, and equipment for major surgeries [29]. Many obstacles limit adequate implementation of health services at the PHC level in South Africa, including the HIV/AIDs pandemic, shortages of HCWs, unequal distribution of resources, and the legacy of the apartheid era [30]. This study was conducted at PHCs.

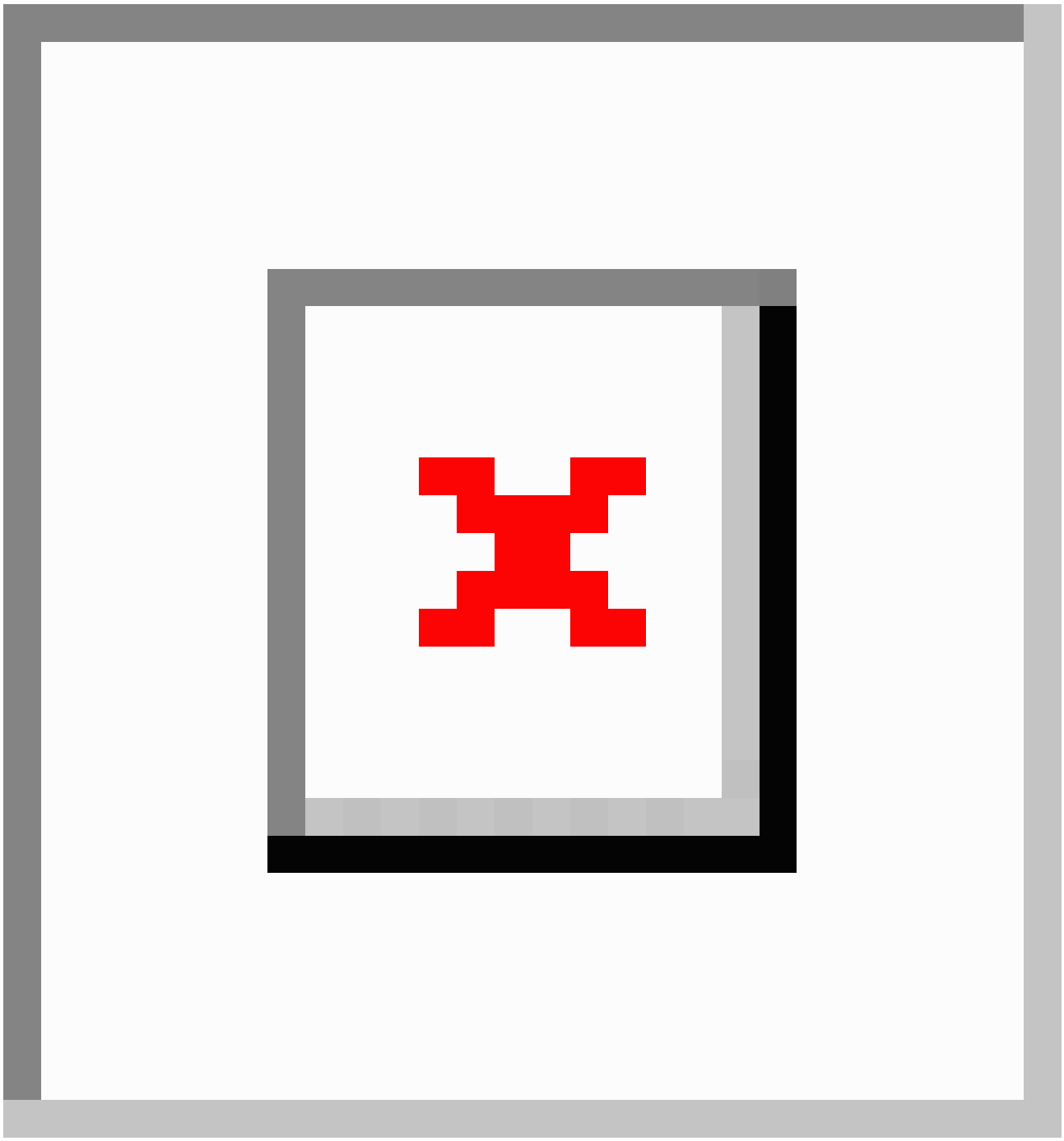
HCD Approach

HCD is based on using "techniques which communicate, interact, empathize and stimulate the people involved, obtaining an understanding of their needs, desires and experiences which

XSL•FO
RenderX

often transcends that which the people themselves actually realized" [31]. This approach encourages stakeholders, experts, and end users—in our case, HCWs—to generate knowledge collaboratively to co-design a medical device [32]. Through involving end users in the design process, HCD allows for the development of devices that are locally and contextually appropriate and can meet the specific health care needs of the population [18]. The key principles of HCD are the active involvement of users and a clear understanding of user and task requirements; iteration of design solutions, where end users provide feedback on design solutions starting early in the process; and making use of multidisciplinary design teams [33]. The HCD approach consists of three iterative stages: (1) inspiration, (2) ideation and prototyping, and (3) formal testing. In this study, we report activities conducted in stage 2, the ideation and prototyping of the *Phefumla* smartphone oximeter development (Figure 2). This builds on our stage-1 findings that explored HCWs' current experience with pediatric pulse oximetry, which will be published elsewhere.

Figure 2. The human-centered design approach.



Participatory Workshops

We conducted participatory workshops, consisting of 4 activities, to facilitate the process of having HCWs co-design a smartphone-based reflectance oximeter (Table 1). The discussion guide is available in Multimedia Appendix 1.

XSL•FO
RenderX

Table . Summary of participatory design workshop activities.

Activities	Resources and tools	Description
1. Design preference	Three 3D-printed plastic prototypes of potential device designs	Participants were given 3 different rapid prototypes and asked questions about the devices' design. These included how they felt about the placement of the sensor on the device; how confident they felt placing the device on a child for a reading; how easy they believe the device would be to clean; and how durable the device was. Participants were encouraged to give suggestions and state preferences.
2. Device use	Two dolls (1 infant sized and 1 slightly larger sized)	Participants were asked to demonstrate how they would use the 3 prototypes on a child, specifically where they would place the sensor for taking a measurement on 2 different sized dolls to represent younger and older infants.
3. Feature ranking	A pile of cards with design features written on them	Participants were asked to rank 11 design features, which were deemed important from stage 1 of the HCD[a] process (eg, battery life), from most to least important. They were also given an opportunity to add their own features on blank cards with markers.
4. User interface	A magnetic board and magnets of different design features of the interface (SpO_2[b] reading, waveform, pulse, bouncing bar, buttons, charging symbols, date, and time)	Participants were asked to arrange interface components as they would like the screen of the device to look. They were provided with a whiteboard marker to include any other features.

[a]HCD: human-centered design.

[b]SpO_2: oxyhemoglobin saturation.

We conducted a pilot design workshop with 3 research nurses from the Desmond Tutu TB Centre to check the quality and coherency of the planned activities. Following the relevant consenting procedures, 1-hour-long design workshops were held with 7 small groups of our sample's HCWs (2-4 HCWs per group). These workshops were conducted by 2 female postgraduate research assistants with comprehensive knowledge and experience of qualitative data collection (EII and LNJ) under the supervision and with the assistance of a pediatric pulmonologist (EDM). Small groups were chosen for pragmatic reasons to minimize disruption to clinical service at the facilities and were conducted in the clinics.

Sampling and Participants

Participants were sampled from a larger pool of participants who had taken part in the previous stage of the study. Stage 1 of the HCD process (inspiration) involved small group discussions with HCWs focusing on barriers and challenges to routine pediatric oximetry use. These HCWs were therefore primed before the co-design workshops to think about the pros and cons of pulse oximeter features. Five clinics in the East subdistrict of Khayelitsha were eligible, and HCWs were purposefully sampled (rich case) using the following inclusion criteria: (1) having experience taking pulse oximeter measurements in children and (2) having taken part in the previous stage of the *Phefumla* study. Participants who had consented in stage 1 to be contacted were followed up to setup face-to-face meetings. Participants were given a small monetary voucher (worth approximately US $15) and provided with refreshments as reimbursement for their time.

Data Collection

Data were collected via audio recordings and photographs taken of activity end results, as well as through comprehensive observation notes. A semistructured workshop guide was developed and used in English, the predominant working language in health care settings in South Africa. However, most participants were native Xhosa speakers, and some discussions were held in Xhosa; LNJ is a native Xhosa speaker and acted as a translator for these sections. The 2 researchers who facilitated the workshops alternated between (1) asking questions and leading facilitation and (2) keeping detailed field notes.

Analysis

Data were analyzed using the framework of exploratory qualitative analysis. Exploratory research is concerned with exploring a phenomenon more deeply to gain a granular understanding of it and has 2 key aspects: open-mindedness and flexibility [34]. Recordings were repeatedly listened to by the 2 researchers who conducted the workshops (EII and LNJ), alongside looking through the captured pictures and written field notes. The wider research team had preidentified key design features of interest based on a rapid scoping review, stage-1 small group discussions, and team expertise. Quotes and notes taken during the workshops were mapped together by EII and LNJ under these categories of design features, using Microsoft Excel. This initial mapping framework was shared and discussed with the entire research team, where findings were discussed and probed. This was done iteratively until the team decided on actionable feedback for the pulse oximeter prototype and shared them with the engineer (MB).

XSL•FO
RenderX

Ethical Considerations

Approval was obtained from the City of Cape Town to recruit HCWs and conduct the project at 7 clinics in the East subdistrict of Khayelitsha. Institutional approval was obtained from Johns Hopkins University (IRB00294436), Stellenbosch University (N22/01/009), and the Swedish Ethics Board (Dnr 2022-01897-01). Facility managers and other relevant gatekeepers were approached after receiving approval to ask for permission to access HCWs. All participants provided written, informed consent. Field notes were anonymized and did not record any identifiable data from HCWs, and recordings were stored in secure local servers to safeguard participant information.

Results

Overview

A total of 7 workshops were conducted with 19 HCWs. The most common reasons for participants from stage 1 not taking part in these stage-2 workshop were being ill, on leave, or absent at the clinic on the scheduled days; having been rotated to a different PHC; or having resigned. All 19 participants retained were nursing staff (including a range of nursing cadres), with 18 (95%) female participants and 1 (5%) male participant.

Activity 1: Selecting a Preferred Prototype

Three 3D-printed prototypes were presented to the groups (Figure 1). The strongest preference was shown for prototype 2 with the sensor in the middle, with 4 (57%) of the 7 groups reaching a consensus on preferring this design. However, this was not unanimous, with 1 (14%) group preferring prototype 1, one (14%) group preferring prototype 3, and 1 (14%) group wanting a combination of prototype 3's larger sensor size with prototype 2's sensor location.

Participants primarily liked prototype 2 because of its sensor location being in the middle, noting that the device would be easier to use on a child as you would not have to angle it to get a reading, it did not matter if you were right- or left-handed, and some participants liked the larger size of the overall device and smaller size of the sensor (compared to the larger sensor of prototype 3). There were some concerns that the device itself was too large and that a smaller device would be easier to use, as well as concerns about having to hold the device without dropping it.

> *It's easier for me to get grasp of the monitor and put it on the child, rather than using the corner.* [HCW, clinic 3]

When discussing the sensor, prototype 3's larger sensor elicited a range of responses, with some HCWs stating that it would be too difficult to use on an infant or young child (eg "it's too big"), whereas others thought the larger sensor size was a benefit, for example:

> *Very much easy [to use] because the sensor is bigger.* [HCW, clinic 1]

> *The sensor is nice and big.* [HCW, clinic 5]

Overall, participants displayed a positive sentiment to this style of device being easy to use on a child, and most participants felt comfortable placing the sensor correctly on a child. Robustness was a concern in several groups, as it was noted that a smartphone screen can break when dropped, and participants offered several modifications in relation to this:

> *It would be better if it were rubber or had a pouch, so it does not break.* [HCW, clinic 1]

> *The back must be rubbery, and the outer part is rubbery too.* [HCW, clinic 3]

> *If it is a glass screen it will [break easily].* [HCW, clinic 4]

All groups stated the device would be very easy to clean, with the most common suggestion for cleaning the device being wiping it with a disinfectant and cloth after each use. All groups felt it would be easy to store as well, with suggestions such as to keep it in a locked drawer, cabinet, or room or to include a storage pouch with the device:

> *Important that it's got a pouch—a bag, so it doesn't get too dusty.* [HCW, clinic 1]

Activity 2: Using the Device on a Child

For activity 2, we asked participants to indicate for each prototype where they would take the pulse oximeter measurements on an infant-sized doll and a larger toddler-sized doll. The purpose was to understand how this novel device would be instinctively applied. The most common location of measurements for infants included the sole of foot, followed by the palm of hand, the hand, the thumb, and toes. These were similar for the older toddler–sized doll, with the sole of foot also being preferred, although HCWs noted that toddlers can kick. Infrequent answers included the wrist, the chest, the forehead, and the neck, which a participant noted would be beneficial as it would not require a child to be undressed. We deliberately did not prompt HCWs to consider specific locations, and it is likely that HCWs defaulted to appendages (ie, hands and feet) that are the most commonly used with a standard pulse oximeter, even if the positioning on those appendages (eg, the palm of the hand) differed.

Activity 3: Design Feature Selection

Table 2 present the results from the feature pile sorting activity, showing the features considered as the 5 most important among the groups. When asked to elaborate on their ranking, participants stated that they first considered what would be essential for the device to function (eg, battery lasting) and that the rest were add-ons (eg, apps installed) that would be nice but were not necessary for core functioning. There was a strong preference displayed for pragmatism in this context:

> *The ones on top are the most important because they're going to sustain the device.* [HCW, clinic 3]

XSL•FO
RenderX

Table . Features ranked among the top 5 for each group.

Feature	Groups (n=7), n (%)	Example quotations for prioritizing features
Portable device	7 (100)	• "You can take it anywhere, for example if it is needed in emergency...then you can take it there." (HCW[a], clinic 4) • "So you can take the sick baby to another room and take the device to the next room" (HCW, clinic 2) • "Because we have three triages in this clinic" (HCW, clinic 3)
Does not break when dropped	6 (86)	• "We are working with kids. It's inevitable that it will fall. It is important that it doesn't break easily when it falls. The kids might not want it and push it away from them and then it falls." (HCW, clinic 4) • "Because we are designing for a small baby not an adult so there's high chances of it falling" (HCW, clinic 1) • "Maybe you're gonna be busy with an emergency so you're going to be scared so you're gonna be shivering or shaking maybe, and the baby will also be fighting you, so at least if it drops it mustn't break easily" (HCW, clinic 2)
Long battery life	6 (86)	• "We're seeing more than 30 children a day and sometimes we don't have time to charge—there's no break when they come. They start to come as early as half past 7 to 4 o'clock so there's no time to say we're still waiting for the battery to get full." (HCW, clinic 4) • "If there's no battery, there's no device." (HCW, clinic 1) • "Loadshedding [of electricity] is happening so it must be charged, and the battery must last" (HCW, clinic 2)
Can measure different parts of the body	5 (71)	• "It doesn't limit you so you can use it on whatever part of the body that you want" (HCW, clinic 2) • "We've got limited sites where you can do accurate readings" (HCW, clinic 1)
Easy to clean	4 (57)	• "Hygiene is very important because we're dealing with kids, so if it's easy to clean then it's more safe." (HCW, clinic 4) • "Because it's used between many patients" (HCW, clinic 5)

[a]HCW: health care worker.

We asked if there were any disagreements in the group, but all groups indicated that they were happy with the consensus reached after discussion.

Participants added the following tangible features: small size, protective cover and storage bag, device holder or stand, and time stamp. However, more generic statements such as "easy to use" and "user friendly" were also added. Some added distractions to the child, such as a colorful screen or pictures. In 1 group, the HCWs also added the inclusion of a probe—the key design feature we were proposing to move away from.

Participants who included "other apps installed" in their overall ranking of design features were asked which apps they desired. The predominant suggestion was to have an app that could also measure temperature and respiratory rate, with a high preference for a multimodal device being displayed. Other app suggestions included an app for referral to hospital emergency wards (known as the *Vula* mobile app in this setting) [35] and an app that referred participants to the emergency medicine practice guidelines [36]. However, there were mixed feelings toward additional apps, and these often were not ranked among the most important features, with the exception of 1 (14%) of the 7 groups. Some of reasons given were that other apps would

XSL•FO
RenderX

not be used; that they would negatively affect the battery life of the device; or that people would use the apps for personal reasons. As some HCWs noted:

> *We're not gonna use other apps.* [HCW, clinic 4]
>
> *I wonder if it's not gonna affect the battery life.* [HCW, clinic 3]
>
> *People overuse it for personal things.* [HCW, clinic 2]

Activity 4: Interface Design

For the user-interface design activity, participants tended to place pulse and SpO_2 readings together (6/7, 86% groups), although there was variability in where on the screen these were placed as well as variability in the size of the icons. Further, 6 (86%) out of the 7 groups included both the waveform as well as the bouncing bar, with the following reasons: if one is not working, the other will; each feature gives you different information; and it makes the device more accessible in the case someone is only familiar with either the bouncing bar or the waveform. The majority (5/7, 71%) of the groups included icons indicating temperature and respiratory rate, further indicating their preference for a multimodal device. Two (29%) groups added a distraction for the child, such as a moving video with sound. The battery was mostly placed at the top of the screen so that HCWs could immediately see whether the device needed to be charged when switched on. When asked what alarms and sounds were wanted, the main preference was for a sound when there was an abnormal reading. Furthermore, the preference was for a loud volume given the noisy environment of the clinics.

Discussion

In this study, we conducted design workshops with South African HCWs to develop a novel, pediatric-specialized pulse oximeter device, to ensure the device is context appropriate. Through the design workshops, we found that HCWs displayed an overall positive and enthusiastic sentiment toward such a device, seeing its value in clinical use with children with hypoxemia. The findings from these workshops were used to select oximeter prototype 2 (Figure 1), with the sensor in the middle, to take forward into the prototype testing stage, with key updates to the robustness and planned user-interface incorporated.

Participants displayed the strongest preference for a device design with a sensor in the middle, feeling that it was overall the easiest to work with. Although participants felt that a smartphone device would be easy to use on a child, clean, and store and felt confident in placing the sensor correctly, some had concerns over the robustness of the device. They provided multiple suggestions to overcome this, such as a pouch, case, or rubber casing, and as a result, we increased the robustness of the device to be able to resist a drop test. However, this may point toward a potential limitation in using a smartphone interface, which HCWs are largely familiar with and have likely had experiences of breakages. This prompted a discussion within the study team on whether the smartphone inside the casing could be replaced with a locally available phone, allowing for a more sustainable repair solution than most traditional medical devices. Although this was not dealt with at this stage of design, the HCW feedback triggered us to reflect on this aspect of the device in more depth and to plan for future prototypes.

We received the least in-depth feedback on the mock placement and use of the prototypes with dolls of infants. One issue may have been the design of the activity, using infant dolls to prompt discussion. As a key challenge in pediatric oximetry, as noted by the HCWs as well, is the children's movement and them becoming agitated with measurements being done, using a real child may have resulted in more reflective responses. The locations that the HCWs largely defaulted to were the thumb, toes, hands, and feet—where oximeter probes are generally used currently, although not with the same versatility. We had hypothesized that a benefit of reflectance oximetry is the range of locations that could be used, which reduces both the HCWs' need to disturb the child and restrict their movement (eg, their forehead or upper back).

Pragmatic concerns arose most strongly during the activity where design features were ranked. These findings speak to the context in which HCWs in LMICs work, where having usable, durable, and long-lasting devices is of the essence—with participants noting that once devices break, they are unlikely to be replaced. This is due to factors such as limited technical and biomedical support and ties into other literature regarding "medical equipment graveyards"—composed of obsolete or otherwise broken biomedical, donated equipment—which are a common occurrence across LMICs [37]. These findings also speak to similar findings in other literature, where opportunities for redesign in pulse oximeters in LMICs included similar themes such as battery charging and durability, probe fit, and sensitivity in pediatric populations [11].

Participants liked the idea of a multimodal device. Although there were various suggestions given for additional design features, a device that could take temperature and respiratory rate readings in addition to SpO_2 was by far the most desirable design feature proposed. This was desirable to participants as one device with multiple modalities is pragmatically beneficial. This speaks to possible opportunities for integration in future device designs and further developments in the field of eHealth; however, this needs to be weighed against risks. There is the risk that more complex devices will be more expensive, have reduced usability, and not be optimized for the oximetry function. Therefore, the benefits need to be weighed against the added value of additional functions, as a device performing a core functionality well could be beneficial over a device that performs poorly across various functionalities [11,38].

Some of the themes raised by our participants were raised in other studies in LMICs, indicating a degree of generalizability. Khayelitsha is considered to be fairly representative of other low-resource, sub-Saharan African settings when considering HIV exposure, tuberculosis mortality rates, and quality of care. However, some contextual factors may be unique to a particular context. For example, Khayelitsha has access to electricity but frequently experiences power supply blackouts (loadshedding), which happens nationwide, meaning that mains-charged devices are acceptable but need a long battery life. In contrast, in other

XSL•FO
RenderX

settings, solar-powered charging was prioritized as access to electricity in health facilities was not universal [11].

Our study has several strengths in being able to rapidly engage with a range of HCWs. However, we also had 3 key limitations. First, we came up with the initial idea for a reflectance pulse oximeter, hypothesizing that this could solve several usability issues for LMIC outpatient settings. Our participants were therefore restricted in their first prototypes to 1 type of oximeter that the research team had chosen. It is possible participants may have preferred an alternative design or traditional transmittance pulse oximeter. It may also have biased our team's presentation of the device and interpretation of the data. However, the workshop researchers had no prior experience in oximetry and led the data collection and analysis process in an attempt to mitigate potential researcher bias. Second, given that our data collection and analysis process were designed to be rapid and pragmatic, we did not extensively pilot the instruments. Lastly, the workshops were conducted primarily in English. Clinical training is done in English and is the language spoken in most professional South African environments. However, it was not the majority of participants' home or first language, which could be a potential limitation. We allowed participants to answer in whatever language they wanted to, and we always had a Xhosa-speaking researcher available to mitigate this limitation.

A contextually appropriate, low-cost, pediatric-specialized, smartphone-based reflectance pulse oximeter was seen to have potential clinical value in the South African context. The process of HCD allowed us to explore HCW's design preferences qualitatively to design a prototype device that would address their specific needs. The overall preference was for a multimodal and pragmatic device, with our rapid participatory approach successfully leading to changes in the oximeter design executed by our engineer.

Acknowledgments

We would like to thank all the health care workers who took part for giving both their time and their wisdom, and the research nurses for their initial feedback when piloting the workshops. The research reported in this publication was supported by the Fogarty International Centre of the National Institutes of Health under award 1R21TW012212-01. In addition, GH receives the financial assistance of the European Union (grant DCI-PANAF/2020/420-028), through the African Research Initiative for Scientific Excellence (ARISE) Pilot Programme. ARISE is implemented by the African Academy of Sciences with support from the European Commission and the African Union Commission. The contents of this document are the sole responsibility of the authors and can under no circumstances be regarded as reflecting the position of the European Union, the African Academy of Sciences, and the African Union Commission. MMvdZ is supported by a career development grant from the EDCTP2 (European and Developing Countries Clinical Trials Partnership) program supported by the European Union (TMA2019SFP-2836 TB lung-FACT2) and the Fogarty International Centre of the National Institutes of Health under award K43TW011028.

Conflicts of Interest

CK and EDM act as independent scientific advisors to the Lifebox Foundation.

Multimedia Appendix 1
Design workshop discussion guide.
[DOCX File, 16 KB - humanfactors_v11i1e54983_app1.docx]

References

1. Lazzerini M, Sonego M, Pellegrin MC. Hypoxaemia as a mortality risk factor in acute lower respiratory infections in children in low and middle-income countries: systematic review and meta-analysis. PLoS One 2015 Sep 15;10(9):e0136166. [doi: 10.1371/journal.pone.0136166] [Medline: 26372640]
2. McCollum ED, King C, Ahmed S, et al. Defining hypoxaemia from pulse oximeter measurements of oxygen saturation in well children at low altitude in Bangladesh: an observational study. BMJ Open Respir Res 2021 Nov;8(1):e001023. [doi: 10.1136/bmjresp-2021-001023] [Medline: 34728475]
3. King C, McCollum ED. Trends in the global burden of paediatric lower respiratory infections. Lancet Infect Dis 2020 Jan;20(1):4-5. [doi: 10.1016/S1473-3099(19)30557-2] [Medline: 31678028]
4. McCollum ED, King C, Colbourn T, et al. Pulse oximetry in paediatric primary care in low-income and middle-income countries. Lancet Respir Med 2019 Dec;7(12):1001-1002. [doi: 10.1016/S2213-2600(19)30358-3] [Medline: 31631030]
5. Rahman AE, Hossain AT, Nair H, et al. Prevalence of hypoxaemia in children with pneumonia in low-income and middle-income countries: a systematic review and meta-analysis. Lancet Glob Health 2022 Mar;10(3):e348-e359. [doi: 10.1016/S2214-109X(21)00586-6] [Medline: 35180418]
6. Lam F, Stegmuller A, Chou VB, Graham HR. Oxygen systems strengthening as an intervention to prevent childhood deaths due to pneumonia in low-resource settings: systematic review, meta-analysis and cost-effectiveness. BMJ Glob Health 2021 Dec;6(12):e007468. [doi: 10.1136/bmjgh-2021-007468] [Medline: 34930758]

XSL•FO
RenderX

7. Badgujar KC, Badgujar AB, Dhangar DV, Badgujar VC. Importance and use of pulse oximeter in COVID-19 pandemic: general factors affecting the sensitivity of pulse oximeter. Indian Chemical Engineer 2020 Nov 27;62(4):374-384. [doi: 10.1080/00194506.2020.1845988]
8. Enoch AJ, English M, Clinical Information Network, McGivern G, Shepperd S. Variability in the use of pulse oximeters with children in Kenyan hospitals: a mixed-methods analysis. PLoS Med 2019 Dec 31;16(12):e1002987. [doi: 10.1371/journal.pmed.1002987] [Medline: 31891572]
9. Ginsburg AS, Tawiah Agyemang C, Ambler G, et al. mPneumonia, an innovation for diagnosing and treating childhood pneumonia in low-resource settings: a feasibility, usability and acceptability study in Ghana. PLoS One 2016 Oct 27;11(10):e0165201. [doi: 10.1371/journal.pone.0165201] [Medline: 27788179]
10. Graham HR, Bakare AA, Gray A, et al. Adoption of paediatric and neonatal pulse oximetry by 12 hospitals in Nigeria: a mixed-methods realist evaluation. BMJ Glob Health 2018 Jun 26;3(3):e000812. [doi: 10.1136/bmjgh-2018-000812] [Medline: 29989086]
11. King C, Boyd N, Walker I, et al. Opportunities and barriers in paediatric pulse oximetry for pneumonia in low-resource clinical settings: a qualitative evaluation from Malawi and Bangladesh. BMJ Open 2018 Jan 30;8(1):e019177. [doi: 10.1136/bmjopen-2017-019177] [Medline: 29382679]
12. Sheikh M, Ahmad H, Ibrahim R, Nisar I, Jehan F. Pulse oximetry: why oxygen saturation is still not a part of standard pediatric guidelines in low-and-middle-income countries (LMICs). Pneumonia (Nathan) 2023 Feb 5;15(1):3. [doi: 10.1186/s41479-023-00108-6] [Medline: 36739442]
13. Zhang J, Tüshaus L, Nuño Martínez N, et al. Data integrity–based methodology and checklist for identifying implementation risks of physiological sensing in mobile health projects: quantitative and qualitative analysis. JMIR Mhealth Uhealth 2018 Dec 14;6(12):e11896. [doi: 10.2196/11896] [Medline: 30552079]
14. Sarin E, Kumar A, Alwadhi V, Saboth P, Kumar H. Experiences with use of a pulse oximeter multimodal device in outpatient management of children with acute respiratory infection during COVID pandemic. J Family Med Prim Care 2021 Feb;10(2):631-635. [doi: 10.4103/jfmpc.jfmpc_1410_20] [Medline: 34041052]
15. Keogh BF, Kopotic RJ. Recent findings in the use of reflectance oximetry: a critical review. Curr Opin Anaesthesiol 2005 Dec;18(6):649-654. [doi: 10.1097/01.aco.0000191893.21444.56] [Medline: 16534307]
16. Fleming KA, Horton S, Wilson ML, et al. The Lancet Commission on diagnostics: transforming access to diagnostics. Lancet 2021 Nov 27;398(10315):1997-2050. [doi: 10.1016/S0140-6736(21)00673-5] [Medline: 34626542]
17. Aranda-Jan CB, Jagtap S, Moultrie J. Towards a framework for holistic contextual design for low-resource settings. International Journal of Design 2016 Dec 31;10(3):43-63 [FREE Full text]
18. Bartlett R, Boyle JA, Simons Smith J, Khan N, Robinson T, Ramaswamy R. Evaluating human-centred design for public health: a case study on developing a healthcare app with refugee communities. Res Involv Engagem 2021 May 30;7(1):32. [doi: 10.1186/s40900-021-00273-2] [Medline: 34053451]
19. Boyd N, King C, Walker IA, et al. Usability testing of a reusable pulse oximeter probe developed for health-care workers caring for children <5 years old in low-resource settings. Am J Trop Med Hyg 2018 Oct;99(4):1096-1104. [doi: 10.4269/ajtmh.18-0016] [Medline: 30141389]
20. Bazzano AN, Martin J, Hicks E, Faughnan M, Murphy L. Human-centred design in global health: a scoping review of applications and contexts. PLoS One 2017 Nov 1;12(11):e0186744. [doi: 10.1371/journal.pone.0186744] [Medline: 29091935]
21. Fisher M, Johansen E. Human-centered design for medical devices and diagnostics in global health. Global Health Innovation 2020 Mar 22;3(1):1-15. [doi: 10.15641/ghi.v3i1.762]
22. Taylor Salisbury T, Atmore KH, Nhambongo I, et al. Integrating human-centred design into the development of an intervention to improve the mental wellbeing of young women in the perinatal period: the Catalyst project. BMC Pregnancy Childbirth 2021 Mar 5;21(1):183. [doi: 10.1186/s12884-021-03675-y] [Medline: 33673826]
23. Asiedu A, Nelson AR, Gomez PP, Tappis H, Effah F, Allen C. “It builds your confidence… you've done well”: healthcare workers' experiences of participating in a low-dose, high-frequency training to improve newborn survival on the day of birth in Ghana. Gates Open Res 2019 May 22;3:1470. [doi: 10.12688/gatesopenres.12936.1] [Medline: 31410394]
24. Triplett NS, Munson S, Mbwayo A, et al. Applying human-centered design to maximize acceptability, feasibility, and usability of mobile technology supervision in Kenya: a mixed methods pilot study protocol. Implement Sci Commun 2021 Jan 7;2(1):2. [doi: 10.1186/s43058-020-00102-9] [Medline: 33413688]
25. Liu C, Lee JH, Gupta AJ, et al. Cost-effectiveness analysis of human-centred design for global health interventions: a quantitative framework. BMJ Glob Health 2022 Mar;7(3):e007912. [doi: 10.1136/bmjgh-2021-007912] [Medline: 35346954]
26. Population data. Western Cape Government. 2020. URL: https://www.westerncape.gov.za/assets/departments/health/h_102_2020_covid-19_population_data.pdf [accessed 2024-05-17]
27. City of Cape Town - 2011 Census suburb Khayelitsha. City of Cape Town. 2013 Jul. URL: https://resource.capetown.gov.za/documentcentre/Documents/Maps%20and%20statistics/2011_Census_CT_Suburb_Khayelitsha_Profile.pdf [accessed 2024-05-17]

XSL•FO
RenderX

28. World Health Organization, Alliance for Health Policy and Systems Research. Primary Health Care Systems (PRIMASYS) : case study from South Africa. World Health Organization. 2017. URL: https://iris.who.int/handle/10665/341145 [accessed 2024-05-17]
29. Lim J, van Loggerenberg E, Chater R. Country profile: South Africa. Bertha Centre for Social Innovation & Entrepreneurship. 2016. URL: https://socialinnovationinhealth.org/2016/12/01/country-profile-south-africa/ [accessed 2024-05-20]
30. Kautzky K, Tollman SM. A perspective on primary health care in South Africa. In: Barron P, Roma-Reardon J, editors. South African Health Review 2008: Health Systems Trust; 2008:17-30 URL: https://www.hst.org.za/publications/South%20African%20Health%20Reviews/sahr2008.pdf [accessed 2024-05-17]
31. Giacomin J. What is human centred design? Des J (Aldershot) 2014 Dec;17(4):606-623. [doi: 10.2752/175630614X14056185480186]
32. Sandholdt CT, Cunningham J, Westendorp RGJ, Kristiansen M. Towards inclusive healthcare delivery: potentials and challenges of human-centred design in health innovation processes to increase healthy aging. Int J Environ Res Public Health 2020 Jun 24;17(12):4551. [doi: 10.3390/ijerph17124551] [Medline: 32599806]
33. Maguire M. Methods to support human-centred design. Int J Hum Comput Stud 2001 Oct;55(4):587-634. [doi: 10.1006/ijhc.2001.0503]
34. Stebbins RA. Exploratory Research in the Social Sciences: SAGE Publications, Inc; 2001. [doi: 10.4135/9781412984249]
35. Secure medical chat & patient referrals. Vula mobile. URL: https://www.vulamobile.com/ [accessed 2023-10-24]
36. Practice guidelines. Emercency Medicine Society South Africa. URL: https://emssa.org.za/practice-guidelines/ [accessed 2023-10-24]
37. Marks IH, Thomas H, Bakhet M, Fitzgerald E. Medical equipment donation in low-resource settings: a review of the literature and guidelines for surgery and anaesthesia in low-income and middle-income countries. BMJ Glob Health 2019 Sep 29;4(5):e001785. [doi: 10.1136/bmjgh-2019-001785] [Medline: 31637029]
38. Baker K, Ward C, Maurel A, et al. Usability and acceptability of a multimodal respiratory rate and pulse oximeter device in case management of children with symptoms of pneumonia: a cross-sectional study in Ethiopia. Acta Paediatr 2021 May;110(5):1620-1632. [doi: 10.1111/apa.15682] [Medline: 33220086]

Abbreviations

HCD: human-centered design
HCW: health care worker
LMIC: low- and middle-income country
PHC: primary health care
SpO_2: oxyhemoglobin saturation

Edited by C Jacob; submitted 29.11.23; peer-reviewed by D Dunsmuir, K Bochtler; revised version received 23.04.24; accepted 01.05.24; published 30.05.24.

Please cite as:
Ilhan EI, Jola LN, van der Zalm MM, Bernstein M, Goussard P, Redfern A, Hesseling AC, Hoddinott G, McCollum ED, King C Designing a Smartphone-Based Pulse Oximeter for Children in South Africa (Phefumla Project): Qualitative Analysis of Human-Centered Design Workshops With Health Care Workers
JMIR Hum Factors 2024;11:e54983
URL: https://humanfactors.jmir.org/2024/1/e54983
doi:10.2196/54983

© Elif I Ilhan, Lucia N Jola, Marieke M van der Zalm, Mike Bernstein, Pierre Goussard, Andrew Redfern, Anneke C Hesseling, Graeme Hoddinott, Eric D McCollum, Carina King. Originally published in JMIR Human Factors (https://humanfactors.jmir.org), 30.5.2024. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on https://humanfactors.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

User-Centered Design and Usability of a Culturally Adapted Virtual Survivorship Care App for Chinese Canadian Prostate Cancer Survivors: Qualitative Descriptive Study

Karen Young[1,2], BCom, MSc; Ting Xiong[1,2], BSc, MSc; Rachel Lee[1], BSc, MHI; Ananya Tina Banerjee[3], PhD; Myles Leslie[4], PhD; Wellam Yu Ko[5], PhD; Quynh Pham[1,2,6,7,8], PhD

[1]Centre for Digital Therapeutics, Techna Institute, University Health Network, Toronto, ON, Canada
[2]Institute of Health Policy, Management and Evaluation, Dalla Lana School of Public Health, University of Toronto, Toronto, ON, Canada
[3]Department of Epidemiology, Biostatistics and Occupational Health, McGill University, Montreal, QC, Canada
[4]School of Public Policy, University of Calgary, Calgary, AB, Canada
[5]Men's Health Research Program, University of British Columbia, Vancouver, BC, Canada
[6]Toronto General Hospital Research Institute, University Health Network, Toronto, ON, Canada
[7]Telfer School of Management, University of Ottawa, Ottawa, ON, Canada
[8]School of Public Health Sciences, University of Waterloo, Waterloo, ON, Canada

Corresponding Author:
Quynh Pham, PhD
Centre for Digital Therapeutics
Techna Institute
University Health Network
Toronto General Hospital, R. Fraser Elliott Building, 4th Floor
190 Elizabeth Street
Toronto, ON, M5G 2C4
Canada
Phone: 1 (416) 340 4800 ext 4765
Email: q.pham@uhn.ca

Abstract

Background: Cultural adaptations of digital health innovations are a growing field. However, digital health innovations can increase health inequities. While completing exploratory work for the cultural adaptation of the *Ned* Clinic virtual survivorship app, we identified structural considerations that provided a space to design digitally connected and collective care.

Objective: This study used a community-based participatory research and user-centered design process to develop a cultural adaptation of the *Ned* Clinic app while designing to intervene in structural inequities.

Methods: The design process included primary data collection and qualitative analysis to explore and distill design principles, an iterative design phase with a multidisciplinary team, and a final evaluation phase with participants throughout the design process as a form of member checking and validation.

Results: Participants indicated that they found the final adapted prototype to be acceptable, appropriate, and feasible for their use. The changes made to adapt the prototype were not specifically culturally Chinese. Instead, we identified ways to strengthen connections between the survivor and their providers; improve accessibility to resources; and honor participants' desires for relationality, accountability, and care.

Conclusions: We grounded the use of user-centered design to develop a prototype design that supports the acts of caring through digital technology by identifying and designing to resist structures that create health inequities in the lives of this community of survivors. By designing for collective justice, we can provide accessible, feasible, and relational care with digital health through the application of Indigenous and Black feminist ways of being and knowing.

(JMIR Hum Factors 2024;11:e49353) doi:10.2196/49353

XSL•FO
RenderX

KEYWORDS

digital health; virtual care; digital therapeutics; prostate cancer; cancer survivorship; user-centred design; usability; supportive care; cultural adaptation; Chinese Canadians

Introduction

Digital health has been posited as a pathway to more equitable and holistic care [1,2]. However, the digital divide, or the capacity for digital technology to exacerbate inequities, has been widely described [3]. Its differential impacts on the social determinants of health are known as the digital determinants of health [4]. Recent years have seen an acceleration of digital health innovations (DHIs) such as digital therapeutics into health care systems, which was supercharged by the COVID-19 pandemic and the resulting widespread implementation of telemedicine [4]. One such digital therapeutic is the *Ned* Clinic ("No Evidence of Disease"), which aims to optimize clinical care and patient self-management through virtual asynchronous care delivery for prostate cancer (PCa) survivors [5]. The *Ned* Clinic platforms, including clinician-led (*Specialist Ned*) and nurse-led (*Ned Nurse*) interventions, were developed at the University Health Network in Toronto, Canada, by a consortium of stakeholders [5].

PCa is the most commonly diagnosed nonskin cancer for Canadian male individuals, and most (99%) are estimated to be diagnosed in male individuals aged 50 years and older [6]. Older adults are negatively impacted by the digital divide [7]. Race, a social determinant of health, is also linked to worse survivorship and care outcomes for PCa survivors, most notably for Black male individuals [8]. Asian (generally defined as East Asian and South Asian ethnicity) male individuals have been found to have better survival rates than the median but are more likely to present with advanced PCa, suggesting systemic issues with identifying health issues and obtaining timely appropriate care [9]. These differences carry over into the delivery of follow-up care, as PCa survivors' care needs and access to care are affected by the complex intersection of ethnicity, culture, and other social and structural factors [10,11].

Cultural adaptation is the process of applying changes to existing health interventions based on "surface" (social and behavioral characteristics) and "deep" (worldview, norms, beliefs, and values) cultural structures [12]. As these structures are known to impact beliefs about illness and well-being, the intent is to provide intervention benefits for communities that have experienced health inequities [13]. Culturally adapted DHIs appear to have been most widely reported in the field of mental health; in contrast, cultural adaptations of cancer survivorship apps have not been published, likely owing to the few interventions in this area [2,14]. The frameworks that appear to be most widely used to adapt health interventions were developed by Bernal et al [15], Resnicow et al [16] (an adaptation of the model by Bernal et al [15]), and Barrera and Castro [17].

However, these guidelines and models often use framings of cultural sensitivity and competency (eg, Resnicow et al [16] and Castro et al [18]), continuing to place the burden of change on individuals rather than addressing the upstream structural determinants of health. These framings can serve to "museumize" and problematize identity categories and culture as causes of ill-health, echoing the long-standing use of culture as a scapegoat to fault specific communities for health inequities. Moreover, defining "culture" for such adaptations can be a complex process in Canada, where culture, race, ethnicity, settler colonialism, and white supremacy (ie, the social and structural determinants of health) all create intersectional and differential lived experiences under a putatively shared identity—Canadian [19-21].

This research reports on the second and final phase of a project to design a cultural adaptation of the patient-facing *Ned* Clinic virtual follow-up care app for Chinese Canadian PCa survivors. In phase 1, we completed formative work distilling a set of themes relevant to survivors' user needs for follow-up and virtual care. Following the user-centered design (UCD) framework, we describe the results of the design and formative evaluation of a culturally adapted prototype of the app.

Methods

Study Design

The overall qualitative descriptive study design was structured using the community-based participatory research (CBPR) and UCD frameworks [22-24]. This study was conducted at the University of Toronto between December 2022 and March 2023 during the COVID-19 pandemic. For communities that face barriers to care, it was found that CBPR practices such as our engagement of a key informant and invitations to community members to share their lived experiences through open-ended interviews are appropriate [1,25]. CBPR concepts were applied to meaningfully involve the community (including several authors of this study) and return the results for their benefit. Here, *community* represents a "symbolic totality as well as a practical multiplicity," as the Chinese Canadian community is highly heterogeneous [26]. We view our participants as a coalition of self-identified Chinese Canadian individuals impacted by PCa survivorship to attend to their differences.

The Chinese Canadian community is an immigrant community that exists as a result of settler colonialism. In recognizing this, we redefine "immigrants" as "people with ancestral roots outside of Indigenous lands, who are beholden to Indigenous laws and epistemologies" [27]. This definition led us to apply a relational paradigm to this project and an axiology of relational accountability. It also provided a pathway to apply several multilevel Indigenous and Black feminist theorizations, guiding principles, and tools [27-29]. These included decolonial theory, *Etuaptmumk* (two-eyed seeing), intersectionality, and cultural safety to inform our conceptualization of digital space as intimately related to land [27,30,31]. This approach allowed us to contextualize the place-related experiences of our participants and uncover their desires for relational and culturally safe care [32]. We noted that these desires are not specifically Chinese,

XSL•FO
RenderX

and this presented an opportunity to design for relationally connected digital health.

UCD is a flexible, iterative, and evidence-based 3-step design process framework that consults, involves, and considers the needs of the end user throughout the entire project [23]. Phase 1 of this study encompasses steps 1 and 2; phase 2 encompasses steps 2 and 3. We present this study according to the Consolidated Criteria for Reporting Qualitative Research (COREQ) guidelines [33].

Step 1: Ideation and Concept Generation

To contextualize the potential use of this app, we sought to understand the structures that impact Chinese Canadian PCa survivors' experiences with follow-up care and virtual care. The results of this phenomenologically informed exploratory-descriptive qualitative study are described elsewhere [34]. Based on the findings of this formative research, we synthesized a list of design principles (Table 1), which we then categorized into the cultural adaptation taxonomy created by Spanhel et al [14] to systematically adapt the patient-facing prototype.

Table 1. Summary of design principles for the adaptation of the *Ned Nurse* patient-facing app.

Research finding	Design principle	Taxonomic classification [14][a]
PHI[b] freedom: patients felt that they were expected to track and remember overwhelming amounts of information.	The system should automatically update, store, and provide access to PHI on demand.	Content components: • (9) Goals of treatment • (10) Methods of treatment
Access to personalized education and information: patients felt that they were unable to access information about their care options and disease status.	The system should provide access to personalized and evidence-based information regarding staging, self-management, and treatment options.	Content components: • (9) Goals of treatment • (10) Methods of treatment
Continuity of care: patients desired a connection with their provider and the ability to communicate during times of need.	The system should improve accessibility and continuity of care, as strong care relationships create a sense of safety.	Content components: • (9) Goals of treatment • (10) Methods of treatment
Security: patients expressed suspicion about digital health because they had concerns about surveillance and security.	The system should be architected and built with a high level of security and privacy.	Methodological components: • (12) Functionality
Accessibility: patients wanted to access care in readable and accessible language formats.	The system should provide readable language and accessible language formats.	Content components: • (5) Language translation • (6) Language tailoring
Digital literacy: patients felt comfortable with their device of choice but desired simplicity, form over function, and accessible help and documentation.	The system should prioritize usability, provide straightforward instruction and support, and maintain simple user interface and user experience design.	Methodological components: • (11) Structure • (12) Functionality • (13) Design and aesthetics
Care coordination: patients felt like they were expected to coordinate their care, as communication between specialists, primary care, and other services were fragmented.	The system should coordinate and provide a clear follow-up appointment schedule.	Content components: • (9) Goals of treatment • (10) Methods of treatment
Resources: patients felt unable to access, refused, or unaware of needed resources such as mental health support.	The system should provide accessible pathways to resources, such as psychological support, supportive care, and financial support.	Content components: • (8) Difference in concepts of mental health and its treatment • (9) Goals of treatment • (10) Methods of treatment

[a]The design principles used to adapt the *Ned* Clinic patient app identified here are classified to the corresponding taxonomic components found in Spanhel et al [14].

[b]PHI: personal health information.

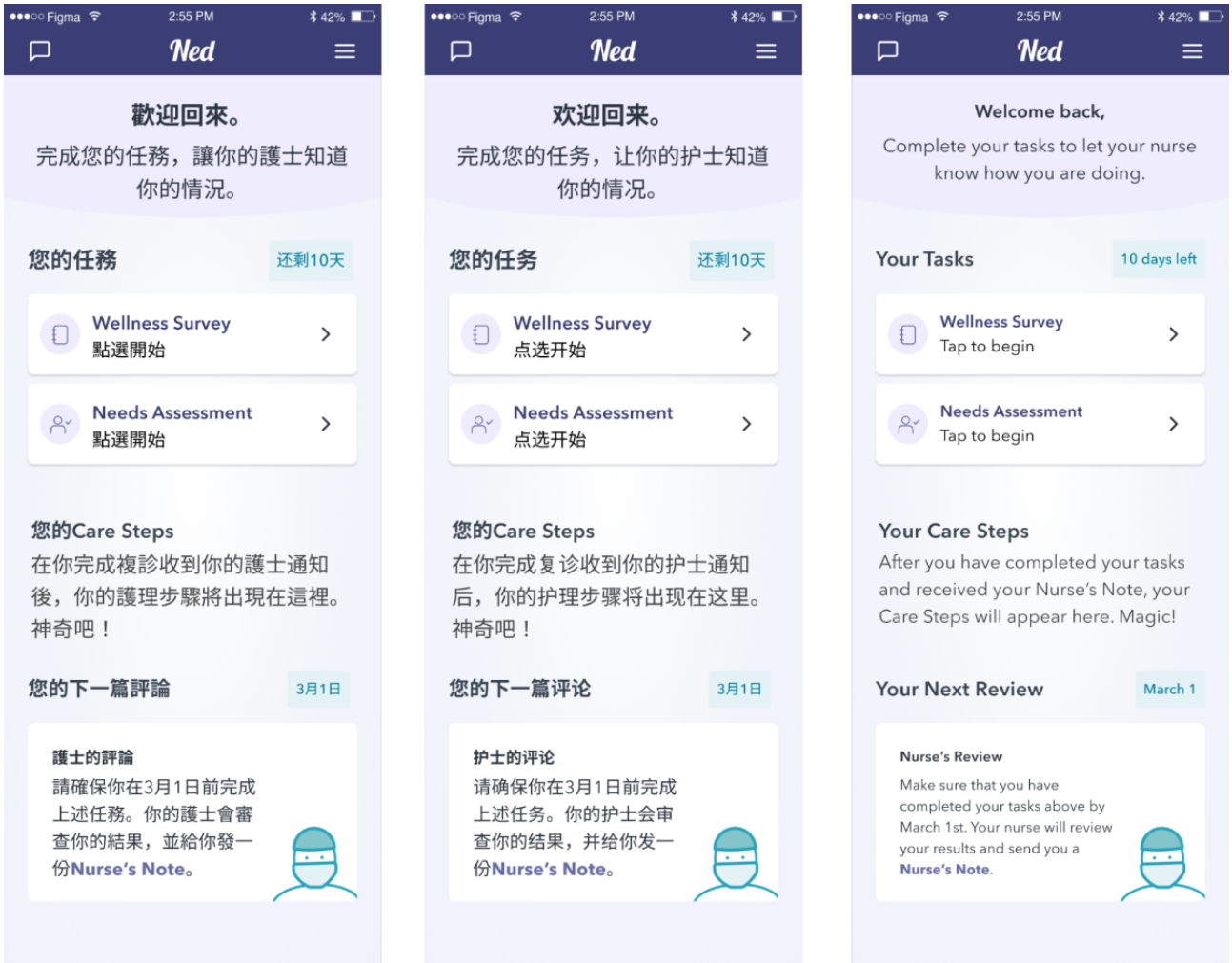
Step 2: Design and Development

We applied these design principles to adapt the *Ned Nurse* patient app for Chinese Canadian survivors. A composite profile of a sample representative user was created to situate the design team during the development of the wireframes. A list of 5 use scenarios was created to guide the adaptation. These scenarios encompassed the design principles created in step 1 and included actions such as completing follow-up tasks, accessing a follow-up care schedule, and using the app to chat with a clinician. All use scenarios are described in the interview guide (Multimedia Appendix 1). Then, the original *Ned Nurse* app wireframes were redesigned to reflect the features required to operationalize these scenarios through the app, resulting in a new prototype. The prototype was created in Figma (Figma Inc.) on an iPhone 13 (Apple Inc) interface. This initial adaptation was iteratively critiqued by a team of researchers and human factors designers to refine the content, user interface, and user experience. Once the adapted prototype was finalized, it was translated from English into written Chinese via the

XSL•FO
RenderX

translation process outlined in Haldane et al [35]. This resulted in 3 versions of the adapted prototype in English, Simplified Chinese, and Traditional Chinese. The app home page in each language version is shown in Figure 1.

Figure 1. Wireframes of the adapted *Ned Nurse* homepage in all 3 language versions.



Step 3: Evaluation

Overview

We empirically evaluated the acceptability, appropriateness, and feasibility of the adaptation through a cultural safety lens [22]. These dimensions are early-stage implementation outcomes and have also been found to be core to the success of DHIs [36,37]. A moderated cognitive walkthrough approach and the think-aloud protocol were used to construct a semistructured interview guide encompassing the 5 scenarios describing usual tasks that an end user might complete through the app [38].

Usability Testing

Facilitators began each test by outlining the usability testing procedure and think-aloud protocol. Context regarding the intended use and deployment of the *Ned Nurse* system was provided. Participants were asked to complete a series of actions for each scenario on the prototype to evaluate its design and functionality. We asked participants to think and speak about improvements they desired during their evaluations. In situations where the participant was unable to access the prototype on their device, they were asked to state their intended actions using the think-aloud protocol to the facilitator, who completed the action in the prototype on their behalf.

Interviews were completed through Microsoft Teams or Zoom (Zoom Technologies Inc). Informed consent for this work was previously obtained as part of overall study consent from participants. Participants were provided with the choice of completing their interview in Cantonese, Mandarin, or English and were also able to choose which language they wished to test the prototype in. The results of each usability test were iteratively analyzed via content analysis. Audio recordings of the participant interviews were translated into English as needed, according to the translation process described previously. A deductive and inductive content analysis approach was used, in which analysis of the data was completed by coders (TX and KY) through a process including open coding, creating categories, and abstraction [39]. Recommendations were applied in real time to create a final prototype that incorporated feedback from each user over the course of usability testing.

Positionality

An important marker of excellent qualitative research is "sincerity" or positionality, which indicates that the researcher has thought about and is reflective and aware about their values,

XSL•FO
RenderX

experiences, biases, and inclinations within their research [40]. Here, the lead researcher reports on their social position, personal experiences, and political and professional beliefs to center the active role that the researcher plays in the framing of the research problem, interpretation of data, methods used, and the reporting of the results [41].

KY is a health informatics trainee and second-generation Chinese Canadian settler who was born and raised in the Greater Vancouver Regional District (GVRD) by a working-class, first-generation immigrant family with roots in southeastern China. She does not have any direct experience with PCa and has not previously provided care for PCa survivors. KY works primarily from a relational paradigm, focusing on the structures, contexts, and relationships that shape the design, development, and implementation of digital therapeutics and health technologies. She led and participated in all study activities.

Setting and Place

This study was conducted in the GVRD, located on the current, unceded, and future territories of the (Tsleil-Waututh, Squamish, and Musqueam) First Nations. The GVRD is home to one of Canada's oldest and largest living Chinese communities, including persons and families whose stories and identities span multiple geographies and generations [42]. The lead (KY) and senior author (QP) established relations with a supportive care program that provides care for Chinese Canadian PCa survivors and a Chinese PCa support group in this area. A key community informant agreed to guide this study and review and approve study materials.

Ethical Considerations

Research ethics approval for this study was obtained from the University of Toronto research ethics board (Human Protocol #43145). Written and verbal informed consent to participate in both phases of the project was obtained from all participants prior to interviews via the REDCap tool (Research Electronic Data Capture; Vanderbilt University), hosted at the University of Toronto. All data collected and disseminated here have been de-identified. Participants were provided with an honorarium of $50.00 CAD ($37.65 USD) per hour in appreciation of their time.

Results

Demographics

Usability testing was performed by 6 user testers, convenience sampled from the pool of 14 survivors and partner-caregivers who participated in the first phase of work as a form of member checking. This sample was also informed by Nielsen-Norman usability testing guidelines [43]. The reasons for nonparticipation were not collected. To protect the privacy of the participants involved in this phase, a demographic overview of the overall research project is provided here. Of the 14 participants in the first phase of this project, all survivors identified as men (n=12, 86%), and all partner-caregivers identified as women (n=2, 14%). A total of 13 (93%) participants indicated that they spoke English as an additional language. Most made an income between CAD $15,000 (US $11,048) and CAD $100,000 (US $73,653; n=12, 86%), lived in an urban area (n=13, 93%), were married (n=12, 86%), were educated beyond high school (n=13, 93%), and were retired (n=9, 64%). A 50/50 split emerged between preferences for smartphone or desktop or laptop use. Most (n=10, 71%) self-rated as being comfortable with their device. Participants indicated that they had 2 or fewer smartphone health apps (n=13, 93%).

Phase 1: Ideation

Table 1 summarizes the user requirement findings that emerged from previous formative research in phase 1 of this project and their subsequent translation to design principles.

Phase 2: Design and Development

Overview of Ned Nurse

An overview of the *Ned Nurse* clinical trial protocol is described by Pham et al [5]. The findings from formative work on the perspectives of health care providers, patients from the wider PCa survivor community, and the service design of the platform are forthcoming. Briefly, *Ned Nurse* digitally operationalizes a nurse-led model of survivorship care. Patients complete a series of tasks or access resources designed to support them in their survivorship. The platform aims to facilitate holistic care for patient quality of life.

Overview of the Adapted Patient-Facing System

The patient-facing adaptation set 2 user-input "care tasks," a validated questionnaire (Expanded Prostate Cancer Index Composite-Clinical Practice [EPIC-CP]) and a needs assessment survey, to constitute a single *Ned Nurse* "review" [5,44]. Language within the app avoided wording such as appointment, visit, and so forth to clarify the differences between synchronous and asynchronous care encounters. The user interface and user experience were designed to draw the user's attention to these tasks on the homepage immediately after login. All features were accessible via an in-app hamburger menu.

User inputs to the questionnaire were triaged via a decision-tree algorithm [45]. The algorithm was designed to return in-app self-management resources within a progress note ("Nurse's Note") automatically available to the user after input submission. If the algorithm detected that the patient required further support, they were prompted to specify domains for follow-up and asked to select their preferred contact method. This action would flag this patient to the nurse for follow-up. Resource links would appear on the homepage after the note was read and cleared.

To ensure that patients were aware of their review schedule, a feature was designed to display the last date, frequency, and next date of their expected reviews. The name of the nurse in charge and an explanation of their *Ned Nurse* role were provided to strengthen the perceived connection between the user and the nurse. This feature also set expectations for manual response times and included a link to users' previous submissions for on-demand access.

Resources were made available in 3 separate categories: symptom self-management advice, PCa information and education, and support and programmatic resources. Within

XSL•FO
RenderX

each category, resources were further categorized. For example, symptom management included resources for symptoms such as anxiety, urinary incontinence, and hot flashes. Each resource provided an overview; relevant self-management steps; off-app links; and the ability to email, print, or save the resource. The feature home page also sectioned resources saved by the user ("Saved Resources") and resources picked for the user ("Picked for Me") by their nurse.

All available and historical prostate-specific antigen and testosterone blood work results were made available in chronological order to the user on-demand in a separate feature. Finally, a chat feature was designed to explore whether users might find it useful. It incorporated both responses in English from an automated support assistant (chatbot) and manually submitted by the nurse. This feature was simulated for evaluation.

Phase 3: Evaluation

Of the 6 participants, 2 (33%) tested in Cantonese, 3 (50%) tested in English, and 1 (17%) tested in Mandarin. These ratios correspond to testing of the Traditional Chinese, English, and Simplified Chinese versions. We note that patients who completed their testing in 1 language were functional to fluent in 1 or all of the other languages and provided critique for multiple versions.

Overall, there was strong agreement that the adaptation presented here would be acceptable, appropriate, and feasible for use, with the exception of the chat feature. Participants agreed that this app would make them feel comfortable and safe by allowing them to have more control over their care, access to resources, and stronger connections to their providers. They were encouraged by its perceived ability to meet their needs by protecting their connection with their providers, leveraging the functional flexibility of digital health, and providing resources beyond what they currently accessed. It was particularly valuable that features could be accessed at their convenience, as some felt that their follow-ups were far too short to meet their needs. Overall, 5 (83%) of 6 participants indicated that the level of support provided by this app was beneficial enough that it should be offered to patients prior to beginning treatment, or even at the point of diagnosis.

Participants' critiques centered on expanding flexibility, access to information, and streamlining responses. They felt that responses for some assessment questions (from 4 to 8 options) were overwhelming and should be reduced (3/6, 50%). English-Chinese translations would increase self-confidence in navigating the health care system. Medication names were spotlighted as particularly difficult. This was noted as an opportunity to expand the app's personal health information (PHI) storage, as a feature containing self-reported PHI (including medications) would be helpful to reference. Pictures and videos were desired instead of textual explanations. Laboratory results were asked to be displayed graphed or with severity indicators by 1 participant, and a text size adjustment function was requested by another.

Support for sexual dysfunction was not requested explicitly but appeared to be implied (3/6, 50%). A sexual therapy resource section was requested by 1 participant. Another noted that they would be more comfortable with nurses gendered as men as they felt uneasy when discussing sexual dysfunction with women. A final participant was keen to indicate that sexual dysfunction was a major area of concern when completing the EPIC-CP questionnaire.

As resources could be accessed on demand, some indicated that more would be beneficial. However, other participants expressed that the number displayed in the prototype were more than sufficient, reflecting our previous study findings on the bifurcated information-seeking behaviors of Chinese Canadian PCa survivors. Participants were also asked if they might find having their imaging results helpful. Although the majority (4/6, 67%) said no, those who said yes (2/6, 33%) were keen on having this information, especially if they needed to travel outside of Canada.

The questionnaire and assessment were generally deemed to be acceptable by most participants (4/6, 33%), with several notable dissents (2/6, 33%). The EPIC-CP question regarding hormonal function was highlighted as confusing by some because the connection between hormonal function and fatigue was not readily apparent. The spiritual domain in the needs assessment was flagged, as some thought that it would not be appropriately addressed by the nurse. Those who felt uncomfortable with this domain noted that they would prefer speaking about these needs to a spiritual leader. Agreement on appropriate response times also varied.

The chat function was deemed possibly helpful but likely unnecessary (4/6, 67%). As all chat interactions were in English, participants who were not confident in their English communication skills felt that their use of this feature would be limited (3/6, 50%). Others felt reminded of troubleshooting cable services rather than feeling connected to their provider. It was emphasized that any opportunity to improve connections to their providers through the app would be appreciated.

Discussion

Principal Findings and Implications

This study provides an applied example of a DHI for Chinese Canadian PCa survivors, which is based on broader principles of collectivism and relationality from Indigenous and Black feminist theory. Our initial aim was to co-design a cultural adaptation of the *Ned* Clinic to provide compassionate care and meet the unmet needs of Chinese Canadian PCa survivors via digital health.

However, attending to cultural adaptation theory and the lived realities of settler colonialism identified gaps to interweave Indigenous and Black feminist teachings. We began by synthesizing design principles that surfaced as critical to our participants and their feelings of comfort and safety when receiving follow-up care. This allowed us to leverage digital health to strengthen relations between the survivor and their providers; improve accessibility to resources; and honor desires for relationality, accountability, and care [46,47]. Rather than adapting by defining Chinese Canadian culture, we co-designed

XSL•FO
RenderX

to intervene in structural causes of health inequities created by settler colonial culture instead [21,48].

We applied *Etuaptmumk* by interweaving strengths from different ways of being and knowing, including those from Indigenous, Western, Chinese, and Black feminist traditions in relation to PCa follow-up and virtual care [27,30,31]. These included prioritizing relational care, accounting for the use of prostate-specific antigen screening as a recurrence monitoring tool, and the benefits of supportive care programs to create adaptation features [30,49]. The EPIC-CP validated questionnaire is a key part of clinical follow-up care, as it allows clinicians to identify possible areas of concern during follow-up [50]. The needs assessment addresses domains beyond clinical care, reflecting the holistic nature of the medicine wheel [51]. Access to resources includes education and guidance for the self-management of concerns across multiple domains. The app presents a "care contract" in the form of a schedule that clearly states the "terms" and dates of the user's follow-ups [52]. It also respects the user's privacy by providing access and allowing them to share their PHI on their terms [53]. Only key inputs are communicated for triage and response. Finally, language access is built into the app as a question of communication accessibility, rather than only culture.

This design approach and these features do not deny the fact that culture is a real influence and can be a source of strength in many peoples' lives. However, we must go beyond implicating culture when designing DHIs for communities made vulnerable and instead address the overarching and underlying structures that create health inequities. Our design approach looked "up" at these structural causes rather than looking "down" and museumizing culture for participants through cultural sensitivity and competency. We demonstrate that a structural approach that applies teachings such as cultural safety and intersectionality can result in DHIs that are found to be acceptable, appropriate, and feasible for use while still leaving room for users to self-define and practice culture on their own terms. We are supporting, not replacing, the *labor* and *acts* of caring with digital health. Beginning with a paradigm shift opened a window to design for collective care, a scalable opportunity to benefit communities beyond Chinese Canadians with this *Ned Nurse* patient-facing app adaptation.

Strengths and Limitations

We have created the first "cultural" adaptation of a PCa follow-up care application for Chinese Canadian survivors. We extended the accessibility of this prototype by offering it in 3 language versions and tested its validity through member checking by returning it to participants who had provided their experiences and expertise as part of the first phase of this project. The findings should be considered with some limitations. Our sample does not fully represent the Chinese Canadian PCa community, as the heterogeneity of the community makes it difficult to recruit a fully representative sample [42]. User testing did not differentiate between results derived from users who interacted with the app themselves and users who directed a facilitator to perform actions on their behalf. However, all participants received the same set of instructions to apply the think-aloud method. A broad description of our theoretical stance, setting and place, methods, and results are provided to enhance understanding. We think of and encourage the transferability of this research as to how it might be made meaningful (ie, valid) for other communities in places where they may be subject to similar constructs and patterns of oppression [32]. Finally, this study does not include the provider perspective, although *Ned* was developed with clinicians who provide follow-up care for patients from this community. Future studies should examine the clinician's perspective on the design and development of similar DHIs, including provision of care through these apps, acceptability and feasibility, and implementation readiness.

Conclusions

This study demonstrates the relationality of Indigenous and Black feminist ontologies, epistemologies, and methodologies to digital health design by providing a worked example of its empirical use for an adaptation of a PCa follow-up care app, the *Ned Nurse* Clinic, for Chinese Canadian PCa survivors. We applied UCD principles to develop a prototype design that supports the relational act of caring through digital technology by identifying structures that create inequities in the experiences of this community of survivors and designing to intervene and provide accessible, connected care instead. We hope that this prototype serves as a tool to help regenerate places of caring, as we have learned from Indigenous and Black feminist scholars' teachings on power, place, and digital technologies.

Acknowledgments

We are deeply grateful for the time and effort gifted to us by the community members who shared their stories. We acknowledge and appreciate the staff and members of the Prostate Cancer Supportive Care Program, including Dr C Higano and Ms M Sundar, and the Richmond Chinese Prostate Cancer Support Group (列治文華人前列腺癌支援網絡), led by Mr D Leung and advised by Dr W Yu Ko. We thank the *Ned* team at the Centre for Digital Therapeutics (University Health Network) for their efforts and commitment to delivering a DHI for Canadian prostate cancer survivors. We give special thanks to Laura Parente for her incisive and human-centered design guidance. The study was funded by the Canadian Institutes of Health Research (CIHR) through the CIHR Canadian Cancer Society Survivorship Grant (CCS: 706713; CIHR: 168606) and the Canada Graduate Scholarship program, as well as by AGE-WELL Networks of Centres of Excellence through the Early Professionals & Inspired Careers in AgeTech (EPIC-AT) Fellowship and Graduate Award. These agencies were not involved in the design or analysis of this study.

XSL•FO
RenderX

Data Availability

The data sets generated and analyzed during this study are available from the corresponding author upon reasonable request.

Authors' Contributions

KY, QP, ATB, ML, and WYK contributed to project conceptualization and study design. KY, RL, and TX contributed to data collection and analysis. KY prepared the first manuscript draft, with contributions from TX, WYK, ATB, ML, and QP. All authors contributed, reviewed, and approved the manuscript.

Conflicts of Interest

QP and the University Health Network (Toronto, Ontario) jointly own intellectual property rights to the *Ned* app. Under the respective agreements with their organizations, QP is entitled to personally benefit from any commercial use of the intellectual property.

Multimedia Appendix 1
Semistructured usability testing interview guide.
[DOCX File , 12 KB - humanfactors_v11i1e49353_app1.docx]

References

1. Pham Q, El-Dassouki N, Lohani R, Jebanesan A, Young K. The future of virtual care for older ethnic adults beyond the COVID-19 pandemic. J Med Internet Res 2022;24(1):e29876 [FREE Full text] [doi: 10.2196/29876] [Medline: 34994707]
2. Pham Q, Hearn J, Gao B, Brown I, Hamilton RJ, Berlin A, et al. Virtual care models for cancer survivorship. NPJ Digit Med 2020;3(1):113 [FREE Full text] [doi: 10.1038/s41746-020-00321-3] [Medline: 32923690]
3. Shaw J, Chandra S, Gordon D, Bickford J, Fujioka J, Yang R, et al. Digital health technologies for more equitable health systems: a discussion paper. Toronto (ON): University of Toronto Joint Centre for Bioethics, Women's College Hospital Institute for Health Systems Solutions and Virtual Care; 2020. URL: https://www.wchwihv.ca/assets/uploads/Digital_Health_and_Equity_Report__September_2020FINAL.pdf [accessed 2023-12-08]
4. Crawford A, Serhal E. Digital health equity and COVID-19: the innovation curve cannot reinforce the social gradient of health. J Med Internet Res 2020;22(6):e19361 [FREE Full text] [doi: 10.2196/19361] [Medline: 32452816]
5. Pham Q, Hearn J, Bender JL, Berlin A, Brown I, Bryant-Lukosius D, et al. Virtual care for prostate cancer survivorship: protocol for an evaluation of a nurse-led algorithm-enhanced virtual clinic implemented at five cancer centres across Canada. BMJ Open 2021;11(4):e045806 [FREE Full text] [doi: 10.1136/bmjopen-2020-045806] [Medline: 33883153]
6. Canadian Cancer Statistics Advisory Committee in collaboration with the Canadian Cancer Society, Statistics Canada and the Public Health Agency of Canada. Canadian Cancer Statistics 2021. Canadian Cancer Society. 2021. URL: http://cancer.ca/Canadian-Cancer-Statistics-2021-EN [accessed 2023-12-08]
7. Latulippe K, Hamel C, Giroux D. Social health inequalities and eHealth: a literature review with qualitative synthesis of theoretical and empirical studies. J Med Internet Res 2017;19(4):e136 [FREE Full text] [doi: 10.2196/jmir.6731] [Medline: 28450271]
8. Bernard B, Muralidhar V, Chen YH, Sridhar SS, Mitchell EP, Pettaway CA, et al. Impact of ethnicity on the outcome of men with metastatic, hormone-sensitive prostate cancer. Cancer 2017;123(9):1536-1544 [FREE Full text] [doi: 10.1002/cncr.30503] [Medline: 28055108]
9. Chao GF, Krishna N, Aizer AA, Dalela D, Hanske J, Li H, et al. Asian Americans and prostate cancer: a nationwide population-based analysis. Urol Oncol 2016;34(5):233.e7-233.e15 [FREE Full text] [doi: 10.1016/j.urolonc.2015.11.013] [Medline: 26725248]
10. Rivas C, Matheson L, Nayoan J, Glaser A, Gavin A, Wright P, et al. Ethnicity and the prostate cancer experience: a qualitative metasynthesis. Psychooncology 2016;25(10):1147-1156 [FREE Full text] [doi: 10.1002/pon.4222] [Medline: 27416079]
11. Wall D, Kristjanson L. Men, culture and hegemonic masculinity: understanding the experience of prostate cancer. Nurs Inq 2005;12(2):87-97. [doi: 10.1111/j.1440-1800.2005.00258.x] [Medline: 15892724]
12. Castro FG, Barrera M, Martinez CR. The cultural adaptation of prevention interventions: resolving tensions between fidelity and fit. Prev Sci 2004;5(1):41-45. [doi: 10.1023/b:prev.0000013980.12412.cd] [Medline: 15058911]
13. Barrera M, Castro FG, Strycker LA, Toobert DJ. Cultural adaptations of behavioral health interventions: a progress report. J Consult Clin Psychol 2013;81(2):196-205 [FREE Full text] [doi: 10.1037/a0027085] [Medline: 22289132]
14. Spanhel K, Balci S, Feldhahn F, Bengel J, Baumeister H, Sander LB. Cultural adaptation of internet- and mobile-based interventions for mental disorders: a systematic review. NPJ Digit Med 2021;4(1):128 [FREE Full text] [doi: 10.1038/s41746-021-00498-1] [Medline: 34433875]
15. Bernal G, Bonilla J, Bellido C. Ecological validity and cultural sensitivity for outcome research: issues for the cultural adaptation and development of psychosocial treatments with Hispanics. J Abnorm Child Psychol 1995;23(1):67-82. [doi: 10.1007/BF01447045] [Medline: 7759675]

XSL•FO
RenderX

16. Resnicow K, Baranowski T, Ahluwalia JS, Braithwaite RL. Cultural sensitivity in public health: defined and demystified. Ethn Dis 1999;9(1):10-21. [Medline: 10355471]
17. Barrera M, Castro FG. A heuristic framework for the cultural adaptation of interventions. Clin Psychol: Sci Pract 2006;13(4):311-316. [doi: 10.1111/j.1468-2850.2006.00043.x]
18. Castro FG, Barrera M, Steiker LKH. Issues and challenges in the design of culturally adapted evidence-based interventions. Annu Rev Clin Psychol 2010;6:213-239 [FREE Full text] [doi: 10.1146/annurev-clinpsy-033109-132032] [Medline: 20192800]
19. Snelgrove C, Dhamoon R, Corntassel J. Unsettling settler colonialism: the discourse and politics of settlers, and solidarity with Indigenous nations. Decolonization 2014;3(2):1-32 [FREE Full text]
20. Harvey M, Neff J, Knight KR, Mukherjee JS, Shamasunder S, Le PV, et al. Structural competency and global health education. Glob Public Health 2022;17(3):341-362. [doi: 10.1080/17441692.2020.1864751] [Medline: 33351721]
21. Czyzewski K. Colonialism as a broader social determinant of health. Int Indig Policy J 2011;2(1):1-16 [FREE Full text] [doi: 10.18584/iipj.2011.2.1.5]
22. Osei-Twum JA, Pulfer E, Banerjee AT. Exploring the applicability of indigenous cultural safety to immigrant health research. In: A Research Agenda for Migration and Health. UK, USA: Edward Elgar Publishing; 2019:138-158.
23. McCurdie T, Taneva S, Casselman M, Yeung M, McDaniel C, Ho W, et al. mHealth consumer apps: the case for user-centered design. Biomed Instrum Technol 2012;Suppl:49-56 [FREE Full text] [doi: 10.2345/0899-8205-46.s2.49] [Medline: 23039777]
24. Sandelowski M. Whatever happened to qualitative description? Res Nurs Health 2000;23(4):334-340. [doi: 10.1002/1098-240x(200008)23:4<334::aid-nur9>3.0.co;2-g] [Medline: 10940958]
25. Pekmezaris R, Kozikowski A, Pascarelli B, Wolf-Klein G, Boye-Codjoe E, Jacome S, et al. A telehealth-delivered pulmonary rehabilitation intervention in underserved hispanic and African American patients with chronic obstructive pulmonary disease: a community-based participatory research approach. JMIR Form Res 2020;4(1):e13197 [FREE Full text] [doi: 10.2196/13197] [Medline: 32012039]
26. Edelman EA. Beyond resilience. Soc Text Duke University Press 2020;38(1):109-130. [doi: 10.1215/01642472-7971127]
27. Tuck E, Wayne YK. Decolonization is not a metaphor. Decolonization 2012;1(1):1-40 [FREE Full text]
28. Wilson S. Research is Ceremony: Indigenous Research Methods. Black Point, Nova Scotia, Canada: Fernwood Publishing Co Ltd; 2008.
29. King TL. The Black Shoals: Offshore Formations of Black and Native Studies. Durham, North Carolina, United States: Duke University Press; 2019.
30. Bartlett C, Marshall M, Marshall A. Two-eyed seeing and other lessons learned within a co-learning journey of bringing together Indigenous and mainstream knowledges and ways of knowing. J Environ Stud Sci 2012;2(4):331-340. [doi: 10.1007/s13412-012-0086-8]
31. Nash JC, Pinto S, editors. The Routledge Companion to Intersectionalities. 1st Edition. Milton Park, Abingdon-on-Thames, Oxfordshire United Kingdom: Taylor & Francis Group; 2022.
32. Tuck E, McKenzie M. Place in Research: Theory, Methodology, and Methods. Milton Park, Abingdon-on-Thames, Oxfordshire, England, UK: Routledge; 2014.
33. Tong A, Sainsbury P, Craig J. Consolidated Criteria for Reporting Qualitative Research (COREQ): a 32-item checklist for interviews and focus groups. Int J Qual Health Care 2007;19(6):349-357 [FREE Full text] [doi: 10.1093/intqhc/mzm042] [Medline: 17872937]
34. Young K, Xiong T, Lee R, Banerjee AT, Leslie M, Ko WY, et al. Honouring the care experiences of Chinese Canadian prostate cancer survivors to cultivate cultural safety and relationality in digital healthxploratory-descriptive qualitative study. J Med Internet Res Preprint posted online on November 28, 2023. [FREE Full text] [doi: 10.2196/49349]
35. Haldane V, Li BP, Ge S, Huang JZ, Huang H, Sadutshang L, et al. Exploring the translation process for multilingual implementation research studies: a collaborative autoethnography. BMJ Glob Health 2022;7(5):e008674 [FREE Full text] [doi: 10.1136/bmjgh-2022-008674] [Medline: 35636804]
36. Proctor E, Silmere H, Raghavan R, Hovmand P, Aarons G, Bunger A, et al. Outcomes for implementation research: conceptual distinctions, measurement challenges, and research agenda. Adm Policy Ment Health 2011;38(2):65-76 [FREE Full text] [doi: 10.1007/s10488-010-0319-7] [Medline: 20957426]
37. Perski O, Short CE. Acceptability of digital health interventions: embracing the complexity. Transl Behav Med 2021;11(7):1473-1480 [FREE Full text] [doi: 10.1093/tbm/ibab048] [Medline: 33963864]
38. Jaspers MWM. A comparison of usability methods for testing interactive health technologies: methodological aspects and empirical evidence. Int J Med Inform 2009;78(5):340-353 [FREE Full text] [doi: 10.1016/j.ijmedinf.2008.10.002] [Medline: 19046928]
39. Elo S, Kyngäs H. The qualitative content analysis process. J Adv Nurs 2008;62(1):107-115. [doi: 10.1111/j.1365-2648.2007.04569.x] [Medline: 18352969]
40. Tracy SJ. Qualitative quality: eight “big-tent” criteria for excellent qualitative research. Qual Inq 2010;16(10):837-851. [doi: 10.1177/1077800410383121]
41. Creswell JW, Poth CN. Qualitative Inquiry and Research Design: Choosing Among Five Approaches. Thousand Oaks, California, United States: SAGE Publications; 2016.

XSL•FO
RenderX

42. Li JTY. The Transcultural Streams of Chinese Canadian Identities. Quebec, Canada: McGill-Queen's Press - MQUP; 2019.
43. Nielsen J. Why you only need to test with 5 users. Nielsen Norman Group. 2000. URL: https://www.nngroup.com/articles/why-you-only-need-to-test-with-5-users/ [accessed 2023-11-26]
44. Chang P, Szymanski KM, Dunn RL, Chipman JJ, Litwin MS, Nguyen PL, et al. Expanded prostate cancer index composite for clinical practice: development and validation of a practical health related quality of life instrument for use in the routine clinical care of patients with prostate cancer. J Urol 2011;186(3):865-872. [doi: 10.1016/j.juro.2011.04.085] [Medline: 21788038]
45. Pfisterer KJ, Lohani R, Janes E, Ng D, Wang D, Bryant-Lukosius D, et al. An actionable expert-system algorithm to support nurse-led cancer survivorship care: algorithm development study. JMIR Cancer 2023;9:e44332 [FREE Full text] [doi: 10.2196/44332] [Medline: 37792435]
46. Arvin M, Tuck E, Morrill A. Decolonizing feminism: challenging connections between settler colonialism and heteropatriarchy. ff 2013;25(1):8-34. [doi: 10.1353/ff.2013.0006]
47. Gaudet JC. Keeoukaywin: the visiting way—fostering an Indigenous research methodology. aps 2019;7(2):47-64 [FREE Full text] [doi: 10.5663/aps.v7i2.29336]
48. Siddiqi A, Hertzman C, Smith BT. The fundamental role of socioeconomic resources for health and health behaviors. In: Fisher EB, Cameron LD, Christensen AJ, Ehlert U, Guo Y, Oldenburg B, et al, editors. Principles and Concepts of Behavioral Medicine: A Global Handbook. New York, NY: Springer New York; 2018:389-413.
49. Roher SIG, Yu Z, Martin DH, Benoit AC. How is Etuaptmumk/Two-Eyed Seeing characterized in Indigenous health research? a scoping review. PLoS One 2021;16(7):e0254612 [FREE Full text] [doi: 10.1371/journal.pone.0254612] [Medline: 34283831]
50. Lee TK, Poon DMC, Ng ACF, Ho T, Singh-Carlson S, Joffres M, et al. Cultural adaptation and validation of the Chinese version of the expanded prostate cancer index composite. Asia Pac J Clin Oncol 2018;14(Suppl 1):10-15. [doi: 10.1111/ajco.12855] [Medline: 29508938]
51. Dapice AN. The medicine wheel. J Transcult Nurs 2006;17(3):251-260. [doi: 10.1177/1043659606288383] [Medline: 16757664]
52. Crosby C, Jakobsen JR. Disability, debility, and caring queerly. Soc Text 2020;38(4):77-103. [doi: 10.1215/01642472-8680454]
53. Ware P, Bartlett SJ, Paré G, Symeonidis I, Tannenbaum C, Bartlett G, et al. Using eHealth technologies: interests, preferences, and concerns of older adults. Interact J Med Res 2017;6(1):e3 [FREE Full text] [doi: 10.2196/ijmr.4447] [Medline: 28336506]

Abbreviations

CBPR: community-based participatory research
COREQ: Consolidated Criteria for Reporting Qualitative Research
DHI: digital health innovation
EPIC-CP: Expanded Prostate Cancer Index Composite-Clinical Practice
GVRD: Greater Vancouver Regional District
PCa: prostate cancer
PHI: personal health information
UCD: user-centered design

Edited by A Kushniruk; submitted 25.05.23; peer-reviewed by J Sharp, X Yan, Y Jain; comments to author 02.08.23; revised version received 05.11.23; accepted 20.11.23; published 01.01.24.

Please cite as:
Young K, Xiong T, Lee R, Banerjee AT, Leslie M, Ko WY, Pham Q
User-Centered Design and Usability of a Culturally Adapted Virtual Survivorship Care App for Chinese Canadian Prostate Cancer Survivors: Qualitative Descriptive Study
JMIR Hum Factors 2024;11:e49353
URL: https://humanfactors.jmir.org/2024/1/e49353
doi:10.2196/49353
PMID:38163295

©Karen Young, Ting Xiong, Rachel Lee, Ananya Tina Banerjee, Myles Leslie, Wellam Yu Ko, Quynh Pham. Originally published in JMIR Human Factors (https://humanfactors.jmir.org), 01.01.2024. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly

XSL•FO
RenderX

cited. The complete bibliographic information, a link to the original publication on https://humanfactors.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Effectiveness and User Perception of an In-Vehicle Voice Warning for Hypoglycemia: Development and Feasibility Trial

Caterina Bérubé[1], PhD; Vera Franziska Lehmann[1,2], PhD; Martin Maritsch[1], PhD; Mathias Kraus[3], Prof Dr; Stefan Feuerriegel[4], Prof Dr; Felix Wortmann[1,5], Prof Dr; Thomas Züger[1,2,6], MD; Christoph Stettler[1,2], Prof Dr; Elgar Fleisch[1,5], Prof Dr; A Baki Kocaballi[7], PhD; Tobias Kowatsch[1,5,8,9], Prof Dr

[1]Centre for Digital Health Interventions, Department of Management, Technology, and Economics, ETH Zurich, Zurich, Switzerland
[2]Department of Diabetes, Endocrinology, Nutritional Medicine and Metabolism, Bern University Hospital, University of Bern, Bern, Switzerland
[3]School of Business, Economics and Society, Friedrich-Alexander-Universität Erlangen-Nürnberg, Nürnberg, Germany
[4]School of Management, Ludwig-Maximilians-Universität München, Munich, Germany
[5]Centre for Digital Health Interventions, Institute of Technology Management, University of St Gallen, St Gallen, Switzerland
[6]Department of Endocrinology and Metabolic Diseases, Kantonsspital Olten, Olten, Switzerland
[7]School of Computer Science, University of Technology Sydney, Sydney, Australia
[8]Institute for Implementation Science in Health Care, University of Zurich, Zurich, Switzerland
[9]School of Medicine, University of St Gallen, St Gallen, Switzerland

Corresponding Author:
Tobias Kowatsch, Prof Dr
Centre for Digital Health Interventions
Department of Management, Technology, and Economics
ETH Zurich
WEV G 216
Weinbergstrasse 56/58
Zurich, 8092
Switzerland
Phone: 41 44 632 94 88
Email: tkowatsch@ethz.ch

Abstract

Background: Hypoglycemia is a frequent and acute complication in type 1 diabetes mellitus (T1DM) and is associated with a higher risk of car mishaps. Currently, hypoglycemia can be detected and signaled through flash glucose monitoring or continuous glucose monitoring devices, which require manual and visual interaction, thereby removing the focus of attention from the driving task. Hypoglycemia causes a decrease in attention, thereby challenging the safety of using such devices behind the wheel. Here, we present an investigation of a hands-free technology—a voice warning that can potentially be delivered via an in-vehicle voice assistant.

Objective: This study aims to investigate the feasibility of an in-vehicle voice warning for hypoglycemia, evaluating both its effectiveness and user perception.

Methods: We designed a voice warning and evaluated it in 3 studies. In all studies, participants received a voice warning while driving. Study 0 (n=10) assessed the feasibility of using a voice warning with healthy participants driving in a simulator. Study 1 (n=18) assessed the voice warning in participants with T1DM. Study 2 (n=20) assessed the voice warning in participants with T1DM undergoing hypoglycemia while driving in a real car. We measured participants' self-reported perception of the voice warning (with a user experience scale in study 0 and with acceptance, alliance, and trust scales in studies 1 and 2) and compliance behavior (whether they stopped the car and reaction time). In addition, we assessed technology affinity and collected the participants' verbal feedback.

Results: Technology affinity was similar across studies and approximately 70% of the maximal value. Perception measure of the voice warning was approximately 62% to 78% in the simulated driving and 34% to 56% in real-world driving. Perception correlated with technology affinity on specific constructs (eg, Affinity for Technology Interaction score and intention to use, optimism and performance expectancy, behavioral intention, Session Alliance Inventory score, innovativeness and hedonic motivation, and negative correlations between discomfort and behavioral intention and discomfort and competence trust; all

XSL•FO
RenderX

P<.05). Compliance was 100% in all studies, whereas reaction time was higher in study 1 (mean 23, SD 5.2 seconds) than in study 0 (mean 12.6, SD 5.7 seconds) and study 2 (mean 14.6, SD 4.3 seconds). Finally, verbal feedback showed that the participants preferred the voice warning to be less verbose and interactive.

Conclusions: This is the first study to investigate the feasibility of an in-vehicle voice warning for hypoglycemia. Drivers find such an implementation useful and effective in a simulated environment, but improvements are needed in the real-world driving context. This study is a kickoff for the use of in-vehicle voice assistants for digital health interventions.

(JMIR Hum Factors 2024;11:e42823) doi:10.2196/42823

KEYWORDS

hypoglycemia; type-1 diabetes mellitus; in-vehicle voice assistant; voice interface; voice warning; digital health intervention; mobile phone

Introduction

Background

Type 1 diabetes mellitus (T1DM) is a chronic condition caused by an inability of the pancreas to produce insulin and requires lifelong insulin therapy [1]. Hypoglycemia, also known as low blood glucose, is a frequent and acute complication in patients with T1DM [2,3]. Symptoms range from autonomic reactions such as trembling, anxiety, and hunger (ie, mild hypoglycemia) to neuroglycopenic reactions such as vision impairment, weakness, or cognitive impairments (ie, severe hypoglycemia) [2,4-6]. Hypoglycemia is a major issue in the context of driving: research has shown that hypoglycemia is associated with a higher risk of car mishaps [7-9]. In fact, drivers experiencing hypoglycemia are recommended by the local authorities [10] to stop the car and treat their condition. However, drivers do not always comply with these recommendations [11,12]. Thus, to help reduce hypoglycemia-related car accidents, there should be an effective warning that informs the driver about an upcoming hypoglycemic episode and supports the driver in coping with the situation. Currently, hypoglycemia can be detected and signaled through flash glucose monitoring (FGM) or continuous glucose monitoring (CGM) devices (ie, wearable receivers connected to a sensor inserted in the subcutaneous tissue of the arm or abdomen) [13]. These allow for glucose monitoring by displaying the values either continuously (ie, CGM) or upon active retrieval (ie, FGM) and deliver alerts in the form of a tone or vibration in case of out-of-range values. However, these devices present limitations in the context of driving. For instance, FGM needs to be held close to the sensor to transfer the data from the subcutaneous sensors to the monitoring device, that is, the driver needs to actively engage in a manual gesture to access the glucose value and to look at a visual display moving the focus of attention from the driving task. In contrast, allowing the drivers to receive an alert in a hands-free mode will facilitate warning reception [14] and lower worry associated with driving with T1DM [15]. However, hypoglycemia is known to cause a decrease in attention [2,4-6], thereby challenging the effectiveness of such devices. As 90% of road accidents are caused by human error, the European Commission has set new safety technologies as mandatory equipment for vehicles as of 2022 (eg, driver drowsiness and distraction warnings and speed assistance) [16]. In-vehicle warning systems for impaired driver states, such as fatigue [17], distraction [18], and breath alcohol concentration [19], are increasingly being developed. However, to the best of our knowledge, there is no existing implementation for hypoglycemia. Such technology would be aligned with the "healing car" concept [20], where vehicles become environments promoting well-being for passengers, including ergonomic seats, ambient lighting, relaxation exercises [21], and detection of health-critical states [22]. This concept is still in its early stages, but it may become a standard in car manufacturing in the future. So far, the only attempts of in-vehicle glucose monitoring are either only proof of concept without user validation [23] or conceptual work [24]. However, the online community clearly expressed a need for in-vehicle glucose monitoring and warning [25].

A growing number of automotive companies are introducing voice assistance technology into their products [26,27]. Voice assistants add value not only for the associated consumer experience but also for their greater safety. Indeed, vocal interactions have been observed to be the least cognitively demanding while driving compared with visual and haptic interactions [28,29]. Moreover, voice assistants are increasingly being implemented to deliver digital health interventions [30-33]. Although research is still in its infancy, efforts have been made to develop voice-based conversational agents to monitor and support individuals with chronic diseases such as cancer, cardiovascular diseases, cognitive disorders, or diabetes [30]. Other recent examples include prevention of excessive alcohol consumption [34], health education and monitoring, physical and mental exercise, and nutrition [35]. Furthermore, a voice assistant delivering a warning is a form of proactive behavior initiated by the computer rather than the user [36,37]. In-vehicle voice assistants can provide personalized and adaptive suggestions, but users may ignore proactive behavior if it is inopportune, violates privacy, or distracts from driving [38-40]. However, emergencies are the most suitable context for proactive behavior that violates privacy [39].

Objectives

Therefore, we investigate the feasibility of an in-vehicle voice warning delivered by a built-in voice assistant to alert and support drivers with T1DM during hypoglycemia. To the best of our knowledge, there have been no investigations on safe and effective in-vehicle hypoglycemia warnings to support drivers with T1DM or on the perception of such technology. Thus, we sought to answer the following research questions (RQs):

XSL•FO
RenderX

- RQ1: How do drivers perceive an in-vehicle voice warning for hypoglycemia while driving?
- RQ2: How effective is an in-vehicle voice warning in prompting drivers to cope with hypoglycemia?

RQ1 refers to the attitude of drivers toward the warning, whereas RQ2 refers to the driver's compliance behavior once the warning is delivered. Answering these RQs will allow us to conclude on the feasibility of an in-vehicle voice warning for hypoglycemia. To control for individual factors influencing the perception of the warning [41], we also assessed technology affinity.

Methods

Study 0: Preliminary Assessment With Healthy Individuals in Simulated Driving

Driving Setting

Participants performed the task in a driving simulator (Carnetsoft Inc) with 3 monitors displaying the front, left, and right views. The central monitor also showed the cockpit and navigation arrows. The participants used a steering wheel and pedals (Logitech Driving Force G29) to control the simulator, which was set to automatic (ie, no clutch or gear shifter). The simulator's computer was connected to a stereo speaker with a subwoofer, which was kept at a constant volume. To control for driving difficulty, 3 environments were used: highway, countryside, and town, with the first and last being the least and most difficult, respectively.

In-Vehicle Voice Warning Simulation

Before testing a hypoglycemia voice warning with people with T1DM, we tested the concept of a car voice assistant as an interface between a dedicated monitoring system and the user with healthy participants. As the participants were not affected by hypoglycemia, the first version of the warning was a simulated low fuel warning ("The car needs a refill. Please pull over and turn off the engine"). Although not health related, it signaled an event of reasonable urgency that required safely stopping the car. Note that the participants were informed that this message aimed to ask them to stop the car as soon as possible and that they did not need to look for a gas station.

The warning was simulated using the Wizard-of-Oz method, where the conversational turns produced by the voice assistant were played by the experimenter [42] from a laptop using predefined keyboard keys. The turns were based on the Google Cloud text-to-speech engine, with a de-DE-Wavenet-C voice, a speed of 1.11 times the normal native speed of the specific voice, and a pitch of −1.20 semitones from the original pitch. The experimenter's computer was connected to the same sound system as the driving simulators so that the voice warning could be heard as part of the driving simulation. No visuals were included.

Voice Warning Evaluation Measures

To assess the RQs, we assessed participants' perception of the warning (self-reported through the modular evaluation of key Components of User Experience [meCUE]; 10 constructs evaluated on a 7-point Likert scale and a general evaluation evaluated on a 10-point scale [43,44]) and participant compliance with the warning (measured by the experimenter manually assessing if the participant would pull over and stop the car following the warning, and reaction time in seconds from the timestamp of the warning to the timestamp of the car fully stopped). As the perception of technology can be influenced by technology-related personality [45], we also measured technology affinity (measured by the Affinity for Technology Interaction [ATI], a 6-point Likert scale [46]). Finally, qualitative feedback was collected informally.

Evaluation Procedure

The participants were welcomed, informed about the procedure, and invited to sit in the simulator. The voice assistant introduced itself and invited the participants to familiarize themselves with the setting, including the 3 environments. The training also screened for motion sickness.

In the experimental session, participants drove 12 times, with 4 blocks of 3 drives each, for approximately 5 minutes per drive. The driving environment's order and starting point varied to minimize habituation. The drive began when the voice assistant prompted participants to start the engine. A timer started to deliver the low fuel warning at either 100 or 200 seconds to add variation and minimize habituation effects. At the end of the session, participants completed the meCUE.

Data Analysis

Participants were characterized by sex, age, and driver's license duration. The ATI was aggregated as a whole, and meCUE items were aggregated per construct. All reports were aggregated across the sample, with mean and SD. Compliance was coded as binary (0=not compliant, 1=compliant) and reported in terms of frequency. Reaction time was aggregated in seconds across participants and phases, with mean and SD.

Study 1: Assessment With Individuals With T1DM in Simulated Driving

Following the iterative approach described earlier, we conducted 3 exploratory iterations. This study was part of a clinical trial registered at ClinicalTrials.gov (NCT04035993).

Driving Setting

The driving setting was the same as in study 0.

In-Vehicle Voice Warning Simulation

On the basis of the results of study 0, we adapted the warning to hypoglycemia instead of low fuel, using the fewest conversational turns possible [47]. To ensure that the drivers were available, the voice assistant started with a receptivity check: "May I disturb you?"

We designed the warning based on the guidelines of the Swiss Diabetes Association [10], which recommends taking carbohydrates and stopping the car as soon as signs of hypoglycemia are noticed. To give the driver a sense of autonomy [48], we designed the warning to suggest eating carbohydrates rather than directly engaging in stopping the car. However, if the driver did not have carbohydrates, they were asked to pull over. On the basis of the feedback, we enhanced

XSL•FO
RenderX

the voice warning used in the following study to recommend pulling over directly (detailed conversation flow is available in Multimedia Appendix 1).

As in study 0, the warning was simulated with a Wizard-of-Oz method [42], and the turns were generated by recording the same voice. However, to reduce fatigue and cognitive load, we decreased the speed and pitch to 0.93 times the normal speed and −4.8 semitones from the original pitch, respectively. As in study 0, the experimenter would play the turns from a Microsoft Windows laptop using predefined keyboard keys to play prerecorded voice sounds. However, in study 1, the laptop program included a visualization mirrored on a smartphone. The visuals consisted of a blue circle that gradually faded in and out when the voice assistant was speaking. As in study 0, the experimenter's computer was connected to the same sound system as the driving simulators, so that the voice assistant could be heard as part of the driving simulation.

Voice Warning Evaluation Measures

Perception assessment focused on evaluating the voice assistant as a trustworthy driving companion. Specifically, participants completed the Acceptance and Use of Technology (AUT) questionnaire [49,50], the Session Alliance Inventory (SAI) [51], and the Emotional Trust and Competence Trust subscales (henceforth Trust) of the Trust and Adoption questionnaire [52].

To assess technology affinity, participants completed the streamlined scale of the Technology Readiness Index (TRI 2.0) [53]. Items were rated on a 5-point Likert scale (ie, 1=totally disagree, 5=totally agree). We also added a question on whether the participants had previous experience with in-vehicle voice assistants (ie, "Have you already had experience with in-vehicle voice assistants?" with a yes or no answer).

Finally, to obtain qualitative and more in-depth feedback for improvement, we conducted a semistructured interview about their experience with the warning (the interview questions are provided in Multimedia Appendix 2).

Evaluation Procedure

The procedure was the same as in study 0, except that participants drove only once for 5 minutes (the evaluation procedure is detailed in Multimedia Appendix 3). Before driving, we ensured that the participants had normal blood glucose levels (5-8 mmol/L).

Data Analysis

The sample of participants was characterized by sex, mean age, and mean duration of their driver's license before the study.

TRI and SAI were aggregated as a whole, and the AUT and Trust items were aggregated per construct. Scores from the negatively formulated questionnaire items were inverted. Previous experience with an in-vehicle voice assistant was reported in terms of frequency. All these reports were aggregated across participants of each iteration, with mean and SD.

To further explain results in perception, they were associated with technology affinity measures. The difference in perception between participants with and without experience with an in-vehicle voice assistant was tested using a 2-sided *t* test, and it was correlated with the TRI constructs using a Pearson test.

Compliance was defined as whether the participant would comply with the warning and was coded as binary (0=did not comply, 1=complied). Reaction time was aggregated in seconds with mean and SD. Compliance behavior was aggregated across participants of each iteration.

Feedback was summarized in positive and negative topics, with a focus on the most prominent suggestions for improvement. Feedback was aggregated across participants of each iteration.

Study 2: Assessment With Individuals With T1DM in Real-World Driving Undergoing Hypoglycemia

Following the iterative approach described earlier, we conducted 2 exploratory iterations. This study was part of a clinical trial registered at ClinicalTrials.gov (NCT04569630).

Driving Setting

Participants drove in Volkswagen Touran on a closed circuit accompanied by a driving instructor. Dual pedals allowed the driving instructor to intervene and stop the car if necessary. The driving environments on the test track corresponded to the environments of the driving simulator used in the previous studies. Straight paths, turns, crossroads, stop signs, and a pedestrian crossing with a doll were used to implement the highway, countryside, and town scenarios. Artificial obstacles (eg, boxes and lines of traffic pylons) were used to simulate the traffic.

In-Vehicle Voice Warning Simulation

On the basis of the participant feedback from study 1, we revised the voice warning and addressed low trust ratings by explaining the cause of the warning. We simulated driving behavior as a trigger to detect hypoglycemia while driving, as in the study by Lehmann et al [54]. We created 2 variations of the simplified hypoglycemia notification—one with a statement of the cause (driving behavior) and one without. The final recommendation was reformulated as stricter but less directive than that in study 1.

In the second iteration, we simplified the conversational flow by removing the receptivity check ("May I disturb you?") and the final recommendation (Multimedia Appendix 1 provides the conversation flow).

We used the Wizard-of-Oz method to simulate the warning, as in studies 0 and 1. We implemented the voice assistant in a smartphone with the same voice as in study 1. However, the experimenter had to control it remotely (outside the car), so we implemented the interaction in a smartphone app controlled by a remote desktop application. The experimenter used the smartphone screen to control the voice warning delivery; therefore, no visualization was included. Because of network-related slowdowns in the remote control, we used a combination of remote control and speech-to-text programing.

Voice Warning Evaluation Measures

All measures were the same as in study 1. Reaction time was calculated from the warning onset until the car reached a velocity of 0. In addition, at the end of the experiment, we

XSL•FO
RenderX

included a questionnaire item asking which of the 2 types of warning they preferred, that is, the warning including a statement of the cause that triggered the warning or the one without it, or if they would not use either of them.

Evaluation Procedure

After welcoming participants and explaining the procedure and simulated voice assistant, the voice assistant introduced itself as an in-vehicle assistant to support drivers with hypoglycemia. The participants then completed a training drive.

The warning was delivered at different stages of hypoglycemia (see the study by Lehmann et al [54]). Drive blocks were defined based on blood glucose levels. In the first phase, the participants drove at normal glucose (5-8 mmol/L). In the second phase, blood glucose level was progressively lowered below the moderate hypoglycemia threshold (3.0 mmol/L) to a target range of 2.0 to 2.5 mmol/L. In the third phase, moderate hypoglycemia was maintained. In the fourth phase, participants drove again with normal blood glucose levels (Multimedia Appendix 3).

To explore the effect of blood glucose level on warning perception and compliance, we delivered a warning at the end of the last drive of each phase. Participants received 2 warnings with an explanation and 2 without, in randomized order.

Data Analysis

Data analysis was carried out as in study 1.

Ethical Considerations

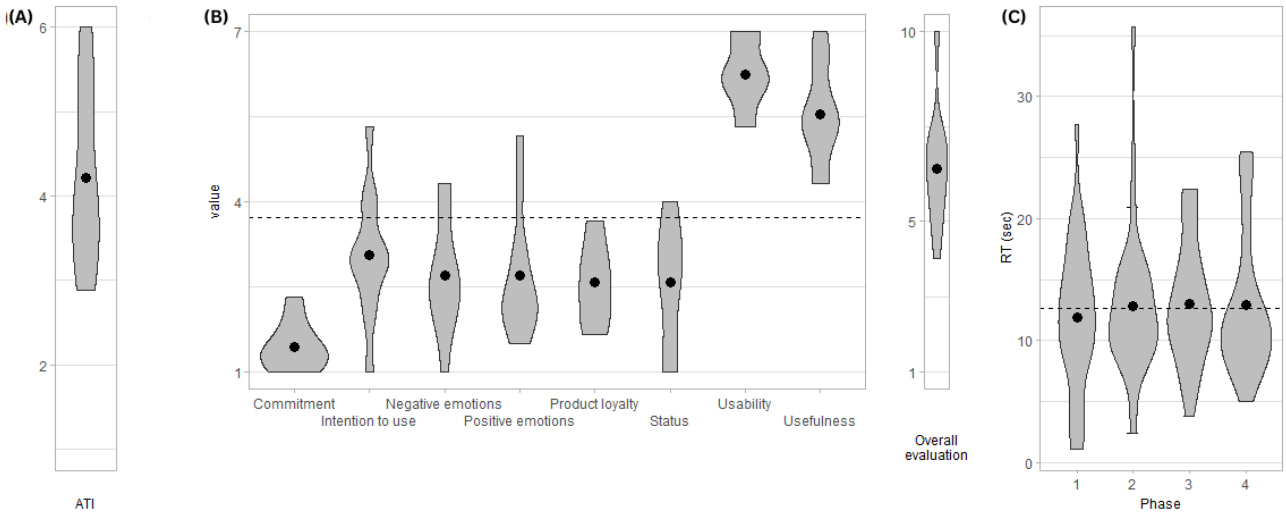
Study 0 was approved by the Ethics Board of ETH Zürich, Switzerland (2019-N-32), and study 1 and study 2 were approved within the context of the HEADWIND study by the cantonal ethics commission of Bern, Switzerland (2020-00685 and 2021-02381, respectively). Study 1 and study 2 are available at ClinicalTrials.gov (NCT04035993 and NCT04569630, respectively). All participants provided written informed consent.

Results

Study 0: Preliminary Assessment With Healthy Individuals in Simulated Driving

Results are summarized in Figure 1.

Figure 1. Violin plots of (A) Affinity for Technology Interaction (ATI; min=1, max=6), (B) score values across the constructs of modular evaluation of key Components of User Experience (meCUE; min=1, max=7, except for overall evaluation, which is min=1, max=10), and (C) reaction time across phases in study 0 (n=11). The dots represent the group mean; the dashed line represents the overall mean. RT: reaction time; sec: seconds.



Recruitment and Participants

We recruited 11 healthy individuals with a valid driver's license via a web advertisement (ie, University of Zurich marketplace). One participant was excluded owing to simulator sickness. Thus, we included 10 participants (n=4, 40% female; n=6, 60% male) with an average age of 30.4 (SD 7.8; range 23-47) years and holding a license for 11 (SD 7.5; range 2-26) years, on average.

Technology Affinity Measure

Participants showed a mean ATI of 4.2 (SD 1; Cronbach α=.91), which is 70% of the maximal value.

Perception Measure

The meCUE (Cronbach α=.7) revealed a mean overall evaluation of 6.4 (SD 1.6), which is 64% of the maximal value. Moreover, the highest mean values were achieved for usability (mean 6.2, SD 0.6, 89%) and usefulness (mean 5.6, SD 0.9, 80%), whereas lower values were observed for commitment (mean 1.5, SD 0.4, 21%), positive emotions (mean 2.7, SD 1.1, 39%), negative emotions (mean 2.7, SD 1, 39%), intention to use (mean 3.1, SD 1.1, 44%), and product loyalty (mean 2.6, SD 0.7, 37%). A low value for negative emotions reflects a more positive evaluation.

To explain the perception results with the technology affinity measure, we correlated each meCUE construct with ATI. We observed a correlation between ATI and intention to use (ρ=0.70; *P*=.02). All the other correlations were not significant at the .05 level.

Compliance Measure

All the participants complied with the warning and stopped the car. Participants took 12.6 (SD 5.7) seconds on average.

XSL•FO
RenderX

Qualitative Feedback

Finally, some participants reported that the voice assistant spoke too fast to deliver information during a driving task without being distracting.

Study 1: Assessment With Individuals With T1DM in Simulated Driving

Results are summarized in Figures 2 and 3.

Figure 2. Violin plots of (A) count of previous experience, (B) score values across the constructs of Technology Readiness Index (TRI; min=1, max=5), (C) score values across the constructs of Acceptance and Use of Technology (AUT; min=1, max=5), (D) Session Alliance Inventory (SAI) scores (min=1, max=5), (E) Trust scores (min=1, max=5), and (F) reaction time across iterations in study 1 (n=18). The dots represent the group means; the dashed line represents the overall mean within an iteration. RT: reaction time; sec: seconds.

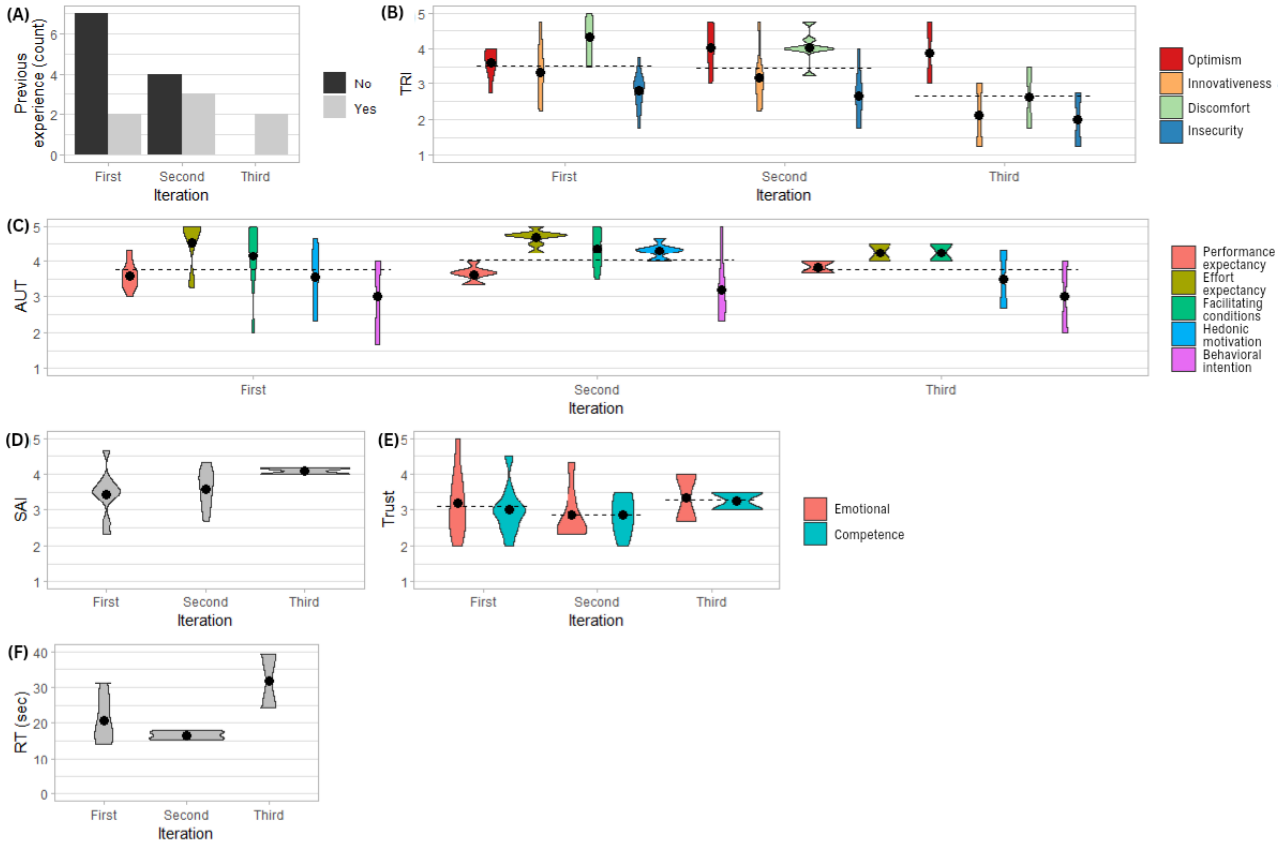
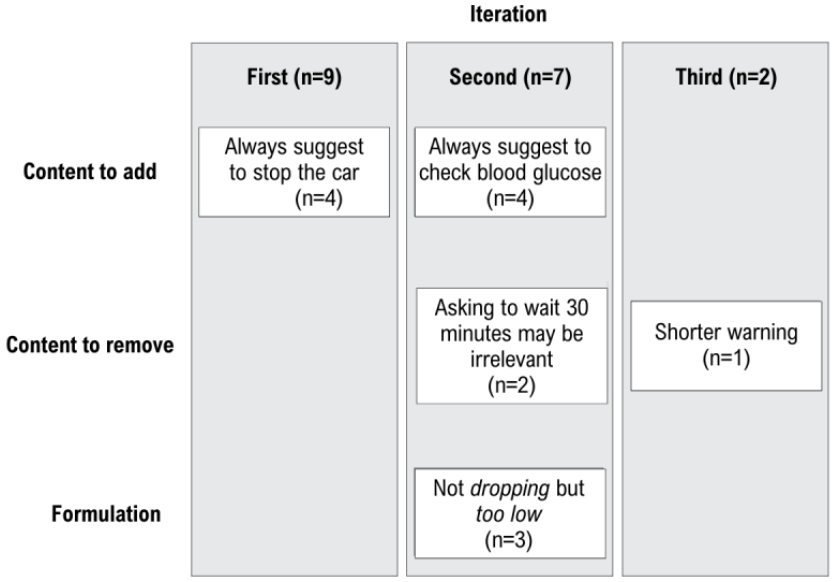


Figure 3. Thematic summary of participants' feedback in study 1 (n=18).



XSL•FO
RenderX

Recruitment and Participants

We recruited 20 patients with T1DM from the Department of Diabetes, Endocrinology, Nutritional Medicine, and Metabolism at the Bern University Hospital. Participants needed functional insulin treatment, good insulin self-management knowledge, a driver's license, and active driving in the past 6 months. We excluded one participant owing to simulator sickness and one participant owing to technical errors in the warning delivery. This resulted in a total of 18 participants (n=6, 33% female and n=12, 67% male; mean age 31.4, SD 7, range 24-44 years; mean driver's license age 13, SD 7.5, range 4.5-28.6 years). The first iteration had 9 participants, the second iteration had 7, and the third iteration had 2 participants. Although the last iteration's sample size was small, it provided useful feedback to improve the warning for study 2.

Technology Affinity Measure

Seven participants had previous experience with an in-vehicle voice assistant (n=2, 28% in the first iteration; n=3, 43% in the second iteration, and n=2, 28% in the third iteration). TRI was 3.4 (SD 0.6, Cronbach α=.85), or 68% of the maximum. Specifically, TRI was 3.5 in the first and second iterations (SD 0.6 and 0.7, respectively) and 2.7 in the third iteration (SD 0.9).

Perception Measure

Perception averaged 3.9 out of 5 (SD 0.8, 78%) and remained stable across iterations. Average AUT (Cronbach α=.81) values were 3.8 (SD 1) in the first iteration, 4 (SD 0.7) in the second iteration, and 3.8 (SD 0.8) in the third iteration. Effort expectancy and facilitating conditions had the highest values across all iterations, whereas behavioral intention always had the lowest values.

SAI (Cronbach α=.79) averaged 3.7 out of 5 (SD 0.5, 74%) and increased slightly over the iterations, from 3.4 (SD 0.7) in the first iteration to 3.6 (SD 0.6) in the second iteration to 4.1 (SD 0.1) in the third iteration.

Trust (Cronbach α=.8) averaged 3.1 out of 5 (SD 0.7, 62%), was stable across constructs, and had the lowest values of the 3 perception measures. Trust averaged 3.1 (SD 0.8) in the first iteration, 2.9 (SD 0.6) in the second iteration, and 3.3 (SD 0.6) in the third iteration.

To explain the perception results with technology affinity, we tested the difference in perception (AUT, SAI, and Trust) between participants with and without previous experience with in-vehicle voice assistants. The means of all constructs, excluding facilitating conditions, were slightly higher for participants with previous experience. However, a 2-sided *t* test revealed no significant result (ie, $P>.05$).

We also correlated perception with TRI and observed a correlation between the optimism construct and performance expectancy (ρ=0.49; $P=.04$), behavioral intention (ρ=0.52; $P=.03$), and SAI (ρ=0.57; $P=.01$). All the other correlations were not significant ($P>.05$).

Compliance Measure

All the participants complied with the warning. In the first iteration, all participants answered *yes* or *no* to the receptivity check ("May I disturb you?") and when asked if they had carbohydrates on hand. Five of the 9 participants answered yes to the latter question, although they did not. Two of those 9 participants stopped the car although they were not explicitly advised to do so. In the second iteration, all participants answered the prompts with yes and stopped the car as advised. One participant gave an affirmative *mhm* when asked, "May I disturb you?" during the hypoglycemic phase but were otherwise compliant. Because we used the Wizard-of-Oz method, the experimenter interpreted the affirmation. However, a current voice assistant might have interpreted it as an error. In the third iteration, both participants answered the prompts with yes and stopped the car. Across iterations, compliance took approximately 22 seconds. In particular, compliance took approximately 20 (mean 20.7, SD 6.2) seconds in the first iteration, approximately 17 (mean 16.7, SD 1.2) seconds in the second iteration, and approximately 31 (mean 31.7, SD 10.6) seconds in the third iteration.

Qualitative Feedback

Participants judged the voice warning as pleasant, simple, and as clear and efficient (n=15, n=11, and n=13, respectively). The topics for improvement are summarized in Figure 3. Note that these results are best understood when compared with Multimedia Appendix 1.

Given that Trust showed the lowest values in the first iteration, in comparison with the other perception measures, we decided to specifically ask participants, in our second and third iterations, what would help them trust the warning more. Of the 9 participants included in both the second and third iterations, 5 (55%) said they would just need to have a prolonged experience with the warning, whereas 3 (33%) said they would need to know what kind of data is used to infer that the driver is about to experience hypoglycemia. One participant did not know what would improve their trust.

Study 2: Assessment With Individuals With T1DM in Real-World Driving Undergoing Hypoglycemia

Results are summarized in Figures 4 and 5.

XSL•FO
RenderX

Figure 4. Violin plots of (A) count of previous experience, (B) score values across the constructs of Technology Readiness Index (TRI; min=1, max=5), (C) score values across the constructs of Acceptance and Use of Technology (AUT; min=1, max=5), (D) Session Alliance Inventory (SAI) scores (min=1, max=5), (E) Trust scores (min=1, max=5), and (F) reaction time across iterations in study 2 (n=20). The dots represent the group means; the dashed line represents the overall mean within an iteration. RT: reaction time; sec: seconds.

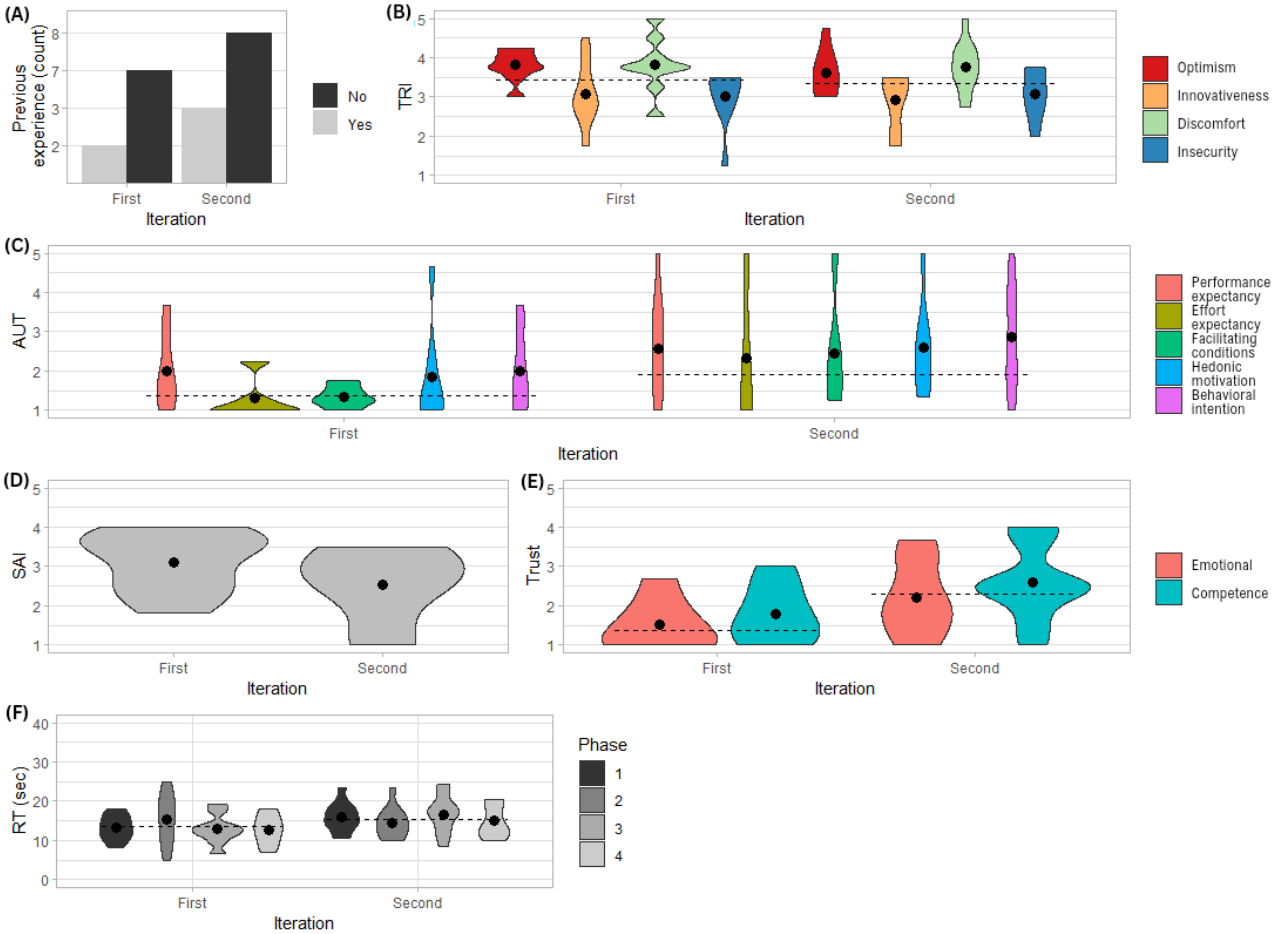
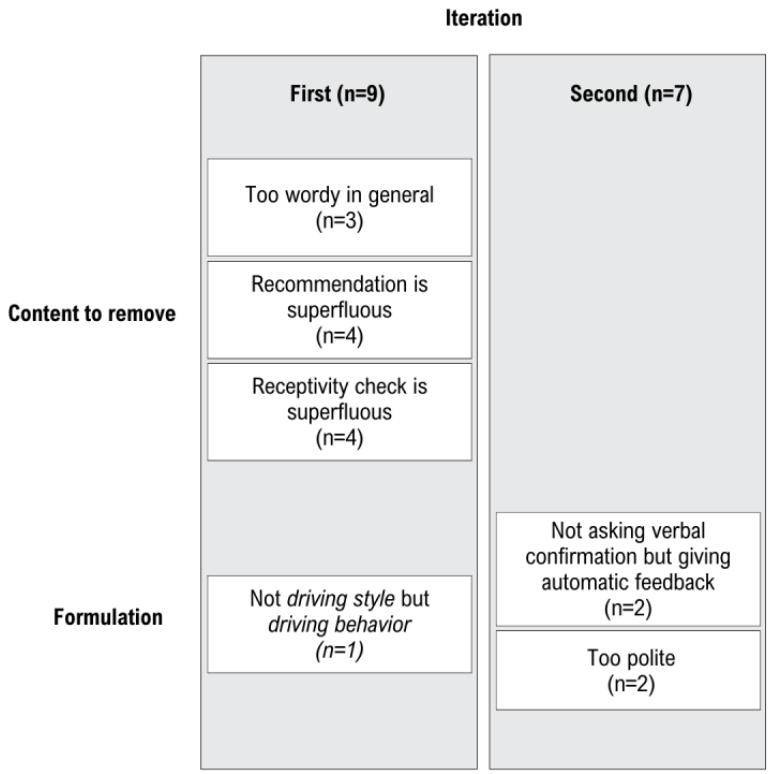


Figure 5. Thematic summary of participants' feedback in study 2 (n=20).



XSL•FO
RenderX

Recruitment and Participants

The recruitment procedure was the same as in study 2. We recruited 21 individuals, and 1 participant was excluded owing to data loss. Thus, we included 20 participants (n=3, 15% female and n=17, 85% male; mean age 40.9, SD 10.6, range 23-57 years; and holding a license on average since 23.7, SD 11.1, range 3.1-42.4 years). The first iteration included a sample of 9 participants and the second iteration included a sample of 11 participants.

Technology Affinity Measure

The pretest measurements revealed that 25% (5/20) of the participants had previous experience with an in-vehicle voice assistant (2 in the first iteration, and 3 in the second iteration), whereas TRI was on average 3.4 (SD 0.7; Cronbach α=.44), which is 68% of the maximal value. In particular, TRI was 3.4 in the first iteration (SD 0.8), and 3.3 in the second iteration (SD 0.7).

Perception Measure

The overall perception score was 1.7 out of 5 (SD 1.3, 34%). The results showed a slight increase in mean AUT (Cronbach α=.95) and Trust (Cronbach α=.85) values between the first and the second iteration, whereas SAI (Cronbach α=.80) showed a slight decrease. AUT also showed a considerable increase in SD. In particular, AUT values were on average 1.4 (SD 1) in the first iteration, and 1.9 (SD 1.6) in the second iteration; SAI was overall 2.8 out of 5 (SD 0.8, 56%). Values were on average 3.1 (SD 0.7) in the first iteration and 2.5 (SD 0.9) in the second iteration; Trust values were on average 1.4 (SD 0.8) in the first iteration, and 2.3 (SD 1.1) in the second iteration.

Similar to study 1, to explain the perception results with the technology affinity measure, we tested the difference in perception (ie, AUT, SAI, and Trust) among participants who had previous experience with in-vehicle voice assistants and those who did not. The means of all perception measures, excluding SAI, were consistently slightly higher in the second iteration. The means of all constructs, excluding SAI, were consistently slightly higher for participants who had previous experience with in-vehicle voice assistants. However, a 2-sided *t* test revealed no significant result (ie, $P>.05$). When correlating each perception measure with TRI, we observed a correlation between innovativeness and hedonic motivation (ρ=0.52; $P=.02$), a negative correlation between discomfort and behavioral intention (ρ=−0.46; $P=.04$), and a negative correlation between discomfort and competence trust (ρ=−0.45; $P=.05$). All the other correlations were not significant ($P>.05$).

Compliance Measure

All the participants complied with the warning. Two drives were excluded: one participant stopped once before the warning was delivered and data from one drive of one participant was lost. The results showed that the reaction time does not seem to vary across glycemic phases and, although minimal, there is a tendency for the reaction time to increase in the second iteration. Participants took 13.6 (SD 4.5) seconds in the first iteration and 15.5 (SD 4.1) seconds in the second iteration.

Preference for the Disclosure of the Triggering Cause

One participant was excluded because of data loss. The results showed that although 10 participants preferred when the warning was delivered with an explanation for the warning being triggered (in this case, driving behavior), 8 participants preferred it without the explanation. One participant stated that they would not use this in-vehicle voice warning either way.

Qualitative Feedback

In general, and similar to study 1, the participants found the communication style pleasant and efficient (n=4 and n=5, respectively). The topics for improvement are summarized in Figure 5. Note that these results are best understood when compared with Multimedia Appendix 1.

Discussion

Principal Findings

Most participants had not previously used an in-vehicle voice assistant, and technology affinity was similar across studies. In general, the voice warning elicited a positive perception, although the perception values were lower in the real-car study. In addition, participants complied with the warning in all studies, and reaction times were shorter in the real-car study than in the simulator study. Finally, the participants preferred the voice warning to be less verbose and prompt fewer interactions with the driver.

Technology Affinity

Although we did not observe a significant effect on the perception of the warning, we suspect that the participants may have experienced a double novelty: using a voice assistant while driving and experiencing a warning from an in-vehicle voice assistant. Thus, future research should include a more balanced sample and compare the perception of a voice assistant–based warning with a standard warning (eg, an acoustic tone). Moreover, although we cannot directly compare ATI (used in study 0) with TRI (used in study 1 and study 2), we can observe that technology affinity was similar across studies. Although ATI showed a mean of 4.2 over 7 (60%), TRI showed a mean of 3.4 over 5 both in study 1 and study 2 (68%). The change in technology affinity measure was the result of an internal discussion between the coauthors, and we recommend the scientific community to use TRI in future research, as it is more widely used and focuses not only on the interaction but also on the general attitudes toward new technologies.

Perception

We observed that AUT, SAI, and Trust values were higher in study 1 (simulated driving) than in study 2 (real-world driving). This evaluation might have been influenced by the driving setting. There can be 2 possible reasons. First, participants may have found the warning to be more distracting in the real car than in the simulator. However, research shows that drivers are more in control in real-world driving than in simulated driving [55]. Second, the technical difficulties in controlling the driver-assistant interaction owing to network slowdowns might have affected the user experience, and thus the perception measures. Future Wizard-of-Oz studies may account for this

XSL•FO
RenderX

methodological weakness with a more accurate text-to-speech technology, avoiding remote control, and reducing interactions.

In addition, TRI seemed to have influenced behavioral intention (AUT) but did not consistently influence the other perception measures (ie, other constructs of AUT, SAI, and Trust). Thus, participants may have been excited about the potential of the voice warning, but they may not have been happy with the actual experience of using it.

Compliance

The reaction times were short enough to ensure a timely reaction to the critical event. Blood glucose can change with a maximum rate of 0.22 mmol/L/min [56]. This means that someone driving with a normal glucose of 5.5 mmol/L might reach hypoglycemia (ie, 3.9 mmol/L) within a minimum of 7.5 minutes. Thus, although experiencing hypoglycemia while driving does not require an abrupt stop but rather a careful pullover maneuver and treating the condition, measuring reaction time provided an insight into the time required to take the first measure (ie, pullover). Interestingly, the reaction time was shorter in the real car (study 2) than in the simulator (study 0 and study 1). This difference may be attributed to the lack of traffic in study 2, which allowed the driver to pull over faster.

Feedback

Although we aimed to keep the warning conversational, participants preferred a more direct notification of the problem without specific recommendations (eg, recommending waiting until the blood glucose is at its normal level) or polite formulations (eg, asking for permission to talk). To the best of the author's knowledge, there was no in-vehicle voice warning at the time of the study, and we mostly relied on the guidelines of the Swiss Diabetes Association [10], while keeping the conversation as simple as possible. The participants' feedback allowed us to improve the warning in this direction.

Implications and Future Directions

Hypoglycemia Warnings

Reportedly, no research has been conducted for in-vehicle applications providing a hypoglycemia warning. However, smartphone apps for hypoglycemic events tracking have been investigated [57]. Although most of the research on glucose monitoring solutions conducted so far focused on diary apps rather than warning delivery, a pilot study on a smartphone-based hypoglycemia warning showed an improved hypoglycemia awareness and a reduction in daytime hypoglycemia [58] (other research is still in the phase of validation [59]). Future research should investigate such outcomes with an in-vehicle extension of this type of application.

In-Vehicle Warnings

Although there seems to be no related work testing the voice assistant of a private vehicle to deliver hypoglycemia warnings, there is a need for "driver-friendly" in-vehicle glucose monitoring solutions, expressed by the online community [25]. In particular, drivers with T1DM have contributed to the Nightscout Foundation [60], a nonprofit organization founded in 2014 and supporting open source technology for T1DM management, with the development of a data-sharing app, able to connect a car to a CGM, and display the glucose trends while driving on the dashboard of the private vehicle [25]. Moreover, there has been conceptual work manifesting the need for collaboration between automotive and medical industries to improve the safety of drivers with T1DM [24]. However, this work has not been followed by any implementation. Furthermore, no testing with the actual users has been conducted. Our work provides preliminary evidence, both in a simulated and a real-world environment.

Needless to say, recognizing hypoglycemia is only one part of glucose monitoring while driving; general imbalance of blood glucose (including hyperglycemia) can be problematic for the driver, if not dangerous [61]. Our work can be extended to hyperglycemia and, therefore, support further the safety of drivers with T1DM.

Finally, using the in-vehicle voice assistant to deliver a warning is compatible with current technology: not only are cars increasingly equipped with voice assistants [26,27] but also the automotive industry is aware of the relevance of using the upcoming "in-car proactivity" [62].

Warning Escalation

Our results showed 100% compliance in all 3 studies. This can only mean that the warning was clear enough for the participant to understand that it was time to pull over. That is, as all studies were run in a controlled setting, where an experimental team was present, and the participant knew they would be recommended to pull over eventually, we can safely assume that the experiments experienced a participant bias [63]. Thus, we cannot conclude that the warning was compelling enough to motivate the participants to comply (see the Limitations section). Nevertheless, the warning should be designed to allow for escalation, whereas in case the driver does not pull over in due time (eg, 2-3 min [56]) or explicitly rejects the warning, delayed reprompts with an increasingly severe tone would be delivered by the voice assistant (eg, "You are at risk of hypoglycemia. Please stop the car safely and check your blood sugar, then risk of hypoglycemia. Pull over now").

Hypoglycemia Detection

Finally, in this paper, we focus on the interface between the hypoglycemia detection system and the driver, with the aim of visually distracting them as little as possible. Although the detection side is beyond the scope of this study, the designed warning is intended to be produced by a voice assistant built into the vehicle. Therefore, how a vehicle monitors blood sugar depends on the technology of the car. For instance, the aforementioned open source app displaying the glucose levels on the dashboard of a private vehicle [25] could be enhanced to connect with the in-vehicle voice assistant and use a voice warning instead of a visual one. Furthermore, research has been conducted on how to detect hypoglycemia from the car's data [54] and from consumer-available wearable devices [64], with the argument that CGM devices can impose a social and financial burden on the individual.

XSL•FO
RenderX

Limitations and Strengths

Despite our best efforts, this research has 3 main limitations.

First, the studies included a relatively small sample size. However, this study includes 3 feasibility studies (ie, a preliminary study with healthy individuals and 2 feasibility studies with individuals with T1DM), and the research presented in this paper is intended to be understood as an iterative development of a hypoglycemic warning. As such, this research aimed to pioneer the use of in-vehicle voice assistants for a driver health-related warning, rather than draw conclusions to be generalized to the population with T1DM. Thus, although we included a total sample size of 48 individuals, each feasibility study provides insight into the changes required by the users, and we provide the scientific community with an opening to the design of in-vehicle voice assistant–based health-related warning. Furthermore, previous studies on digital health systems used a similar sample size [65-67]. Thus, we believe that although the sample size does not allow drawing conclusions on the interaction of drivers with T1DM with in-vehicle hypoglycemia warnings, it still reports pioneer research.

Second, the studies were conducted over a short period. The participants had only a short-term experience with the warning. Perception and compliance may therefore be influenced by the novelty of such an experience, whereas perception may stabilize with repeated experience [68]. Future research should investigate the user experience of the warning in a longitudinal study. Third, these studies did not control for all potentially confounding variables related to real-world traffic and driver's priorities. For instance, both simulator and real-car experiments involved disadvantages: while assessing the warning in a simulator allowed a controlled and safe experiment, such a setting remains artificial and lacks external validity. In contrast, while testing it in a real car increased the ecological validity of the human-machine interaction, it did not allow for as much traffic and speed variation as was possible in the simulator. Future research should investigate the effects of real-world traffic on the perception of the warning and compliance behavior. Moreover, receiving a warning in the presence of a team of experimenters may have influenced the participant's verbal and behavioral responses; participants knew they would receive the warning sooner or later and had no reason not to follow it (eg, ignoring the warning because of being late for an appointment). In a real situation, drivers may not respond as expected or may even ignore the warning. Future research should test such a warning in a more ecological context, for instance, in a field study where the driver may not fall for a participant bias [69].

Finally, as we aimed to test a voice warning, our studies used a Wizard-of-Oz methodology to avoid problems related to natural language processing. Note that our studies were conducted in German-speaking Switzerland, where the German accents easily vary from region to region. As this aspect was beyond the scope of our research, we did not implement a working voice assistant or account for potential fallback intents triggered by the voice assistant's failure to understand the user. Future research should push this research further and examine the potential danger of delayed treatment of hypoglycemia owing to the voice assistant's natural language processing errors.

Conclusions

Although hypoglycemia increases the risk of car mishaps [7,8], current solutions (eg, CGM and FGM) require visual human-machine interaction, which is inappropriate for an in-vehicle context. As voice assistants are increasingly present in private vehicles [26,27] and the European Commission fosters safety technologies inside the car [16], we propose to warn the driver of their critical health state through a voice assistant–based health warning. This paper reports on an iterative development and assessment of a hypoglycemia warning. In particular, we conducted in 3 studies: a preliminary study using a simulator with healthy participants, a test with individuals with T1DM in a simulator, and a test with individuals with T1DM in a real car. This gradual increase in authenticity in the experimental design allowed us to increase the ecological validity of our results while keeping experimental control. To the best of our knowledge, this is the first attempt of such a comprehensive feasibility assessment of an in-vehicle voice warning for hypoglycemia. Our results suggest that a voice warning can be useful, but that proactive behavior in voice assistants is still emerging and unfamiliar. We hope that these preliminary findings will foster future research to further develop in-vehicle hypoglycemia warnings.

Acknowledgments

The authors would like to thank the Swiss Federal Department of Defense, Civil Protection and Sports for providing the closed circuit used in study 2.

This work was partly funded by the Swiss National Science Foundation, Sinergia Project (CRSII5 183569).

Data Availability

The data supporting the findings of this study are available from the last author, TK, upon reasonable request.

Authors' Contributions

CS and EF were responsible for the oversight, leadership, and management of the research activity and for funding acquisition. CB, MM, VFL, MK, SF, TZ, FW, CS, EF, and TK were responsible for the methodology. CB and TK were responsible for the formulation of the research goals and aims, conceptualizing the voice warning. CB and MM were responsible for developing the voice warnings. CB, MM, FW, and TK were responsible for developing the driving scenarios. FW, TK, and TZ were responsible

XSL•FO
RenderX

for providing the driving simulator, the real car, and the closed circuit. CB and MM were responsible for recruiting participants for study 0. VFL, TZ, and CS were responsible for recruiting participants for studies 2 and 3. CB, MM, and VFL were responsible for data collection. CB was responsible for the data analysis and presentation and the first draft of this manuscript. VFL, MK, SF, TZ, CS, EF, ABK, and TK were responsible for critical feedback and final revisions of the manuscript.

Conflicts of Interest

CB, VFL, SF, FW, TZ, CS, EF, and TK are affiliated with the Centre for Digital Health Interventions, a joint initiative of the Department of Management, Technology, and Economics at ETH Zürich, and the Institute of Technology Management at the University of St Gallen, which is funded in part by the Swiss health insurer CSS. EF and TK are also the cofounders of Pathmate Technologies, a university spin-off company that creates and delivers digital clinical pathways. However, neither CSS nor Pathmate Technologies were involved in any way in the design, interpretation, analysis, or writing. All other authors declare no other conflicts of interest.

Multimedia Appendix 1
Original (German) and translated version of the conversation flow of the hypoglycemia voice warning in study 1 and study 2.
[PDF File (Adobe PDF File), 554 KB - humanfactors_v11i1e42823_app1.pdf]

Multimedia Appendix 2
Questions used in the semistructured interview about participants' experience with the warning conducted in study 1 and study 2.
[PDF File (Adobe PDF File), 209 KB - humanfactors_v11i1e42823_app2.pdf]

Multimedia Appendix 3
Illustration of the procedure across studies.
[PDF File (Adobe PDF File), 174 KB - humanfactors_v11i1e42823_app3.pdf]

References

1. Moini J. Type 1 diabetes. Epidemiol Diabetes 2019:75-90. [doi: 10.1016/b978-0-12- 816864-6.00006-7]
2. Cryer PE, Davis SN, Shamoon H. Hypoglycemia in diabetes. Diabetes Care 2003 Jun;26(6):1902-1912 [FREE Full text] [doi: 10.2337/diacare.26.6.1902] [Medline: 12766131]
3. Malkani S, Kotwal A. Frequency and predictors of self-reported hypoglycemia in insulin-treated diabetes. J Diabetes Res 2017;2017:7425925 [FREE Full text] [doi: 10.1155/2017/7425925] [Medline: 28913365]
4. Allen KV, Pickering MJ, Zammitt NN, Hartsuiker RJ, Traxler MJ, Frier BM, et al. Effects of acute hypoglycemia on working memory and language processing in adults with and without type 1 diabetes. Diabetes Care 2015 Jun;38(6):1108-1115 [FREE Full text] [doi: 10.2337/dc14-1657] [Medline: 25758768]
5. Leung E. New and not so new information on hypoglycemia and prevention strategies. Journal: The Official Publication of the Ontario Occupational Health Nurses Association 2018;37(1):20-22 [FREE Full text]
6. Warren RE, Frier BM. Hypoglycaemia and cognitive function. Diabetes Obes Metab 2005 Sep;7(5):493-503 [FREE Full text] [doi: 10.1111/j.1463-1326.2004.00421.x] [Medline: 16050942]
7. Blumer I, Kenshole AB, Stilman J, Lewis GF. Insulin-treated diabetes and driving: legal jeopardy and consequences of hypoglycemia. Can J Diabetes 2019 Apr;43(3):221-223 [FREE Full text] [doi: 10.1016/j.jcjd.2018.08.201] [Medline: 30348593]
8. Cox DJ, Penberthy JK, Zrebiec J, Weinger K, Aikens JE, Frier B, et al. Diabetes and driving mishaps: frequency and correlations from a multinational survey. Diabetes Care 2003 Aug;26(8):2329-2334 [FREE Full text] [doi: 10.2337/diacare.26.8.2329] [Medline: 12882857]
9. Waris A, Jose R, Mathew N, Jacob JJ. Attitudes to driving among patients with diabetes mellitus in Punjab (ADD-Punjab study). Prim Care Diabetes 2022 Feb;16(1):34-40 [FREE Full text] [doi: 10.1016/j.pcd.2021.11.007] [Medline: 34872840]
10. Fankhauser B. Diabetes und autofahren: worauf sie am steuer achten sollten. d-journal. 2018 Nov 2. URL: https://www.d-journal.ch/uebriges-keine/diabetes-und-autofahren-worauf-sie-am-steuer-achten-sollten/ [accessed 2023-12-22]
11. Cox DJ, Gonder-Frederick L, Clarke W. Driving decrements in type I diabetes during moderate hypoglycemia. Diabetes 1993 Feb;42(2):239-243. [doi: 10.2337/diab.42.2.239] [Medline: 8425660]
12. Roberts AJ, Moss A, Malik FS, Taplin C, Pihoker C, Hirsch IB, et al. Driving safety in adolescents and young adults with type 1 diabetes. Diabetes Spectr 2020 Nov;33(4):352-357 [FREE Full text] [doi: 10.2337/ds20-0021] [Medline: 33223774]
13. Cappon G, Vettoretti M, Sparacino G, Facchinetti A. Continuous glucose monitoring sensors for diabetes management: a review of technologies and applications. Diabetes Metab J 2019 Aug;43(4):383-397 [FREE Full text] [doi: 10.4093/dmj.2019.0121] [Medline: 31441246]
14. Landry S, Jeon M, Lautala P, Nelson D. Design and assessment of in-vehicle auditory alerts for highway-rail grade crossings. Transp Res Part F Traffic Psychol Behav 2019 Apr;62:228-245 [FREE Full text] [doi: 10.1016/j.trf.2018.12.024]

XSL•FO
RenderX

15. Böhme P, Bertin E, Cosson E, Chevalier N, GEODE group. Fear of hypoglycaemia in patients with type 1 diabetes: do patients and diabetologists feel the same way? Diabetes Metab 2013 Feb;39(1):63-70 [FREE Full text] [doi: 10.1016/j.diabet.2012.10.006] [Medline: 23266467]
16. Road safety: commission welcomes agreement on new EU rules to help save lives. European Commission. 2019 Mar 26. URL: https://ec.europa.eu/commission/presscorner/detail/en/IP_19_1793 [accessed 2023-12-22]
17. Sikander G, Anwar S. Driver fatigue detection systems: a review. IEEE Trans Intell Transport Syst 2019 Jun;20(6):2339-2352 [FREE Full text] [doi: 10.1109/tits.2018.2868499]
18. Fridman L, Toyoda H, Seaman S, Seppelt B, Angell L, Lee J, et al. What can be predicted from six seconds of driver glances? In: Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems. 2017 Presented at: CHI '17: CHI Conference on Human Factors in Computing Systems; May 6-11, 2017; Denver, CO URL: https://doi.org/10.1145/3025453.3025929 [doi: 10.1145/3025453.3025929]
19. Ljungblad J, Hök B, Allalou A, Pettersson H. Passive in-vehicle driver breath alcohol detection using advanced sensor signal acquisition and fusion. Traffic Inj Prev 2017 May 29;18(sup1):S31-S36 [FREE Full text] [doi: 10.1080/15389588.2017.1312688] [Medline: 28368660]
20. Koch K, Liu S, Berger T, Wortmann F. Towards the healing car: investigating the potential of psychotherapeutic in-vehicle interventions. In: Proceedings of the European Conference on Information Systems. 2020 Presented at: ECIS 2020; June 15-17, 2020; Marrakech, Morocco URL: https://aisel.aisnet.org/ecis2020_rip/8
21. Paredes PE, Zhou Y, Hamdan NA, Balters S, Murnane E, Ju W, et al. Just breathe: in-car interventions for guided slow breathing. Proc ACM Interact Mob Wearable Ubiquitous Technol 2018 Mar 26;2(1):1-23 [FREE Full text] [doi: 10.1145/3191760]
22. Albadawi Y, Takruri M, Awad M. A review of recent developments in driver drowsiness detection systems. Sensors (Basel) 2022 Mar 07;22(5):2069 [FREE Full text] [doi: 10.3390/s22052069] [Medline: 35271215]
23. Merickel J, High R, Smith L, Wichman C, Frankel E, Smits K, et al. Driving safety and real-time glucose monitoring in insulin-dependent diabetes. Int J Automot Eng 2019;10(1):34-40 [FREE Full text] [doi: 10.20485/jsaeijae.10.1_34] [Medline: 34306907]
24. Kerr D, Olateju T. Driving with diabetes in the future: in-vehicle medical monitoring. J Diabetes Sci Technol 2010 Mar 01;4(2):464-469 [FREE Full text] [doi: 10.1177/193229681000400228] [Medline: 20307408]
25. Hoskins M. Glooko's role in revolutionizing diabetes data usability. Healthline. 2022 Apr 7. URL: https://www.healthline.com/diabetesmine/diabetes-data-in-cars-wearenotwaiting [accessed 2023-12-22]
26. In-car voice assistant consumer adoption report 2019. Voicebot.ai. URL: https://voicebot.ai/in-car-voice-assistant-consumer-adoption-report-2019/ [accessed 2023-12-22]
27. In-car voice assistant consumer adoption report 2020. Voicebot.ai. URL: https://voicebot.ai/in-car-voice-assistant-consumer-adoption-report-2020/ [accessed 2023-12-22]
28. Baron A, Green P. Safety and usability of speech interfaces for in-vehicle tasks while driving: a brief literature review. The National Academies of Sciences, Engineering, and Medicine. 2006. URL: https://trid.trb.org/view/806083 [accessed 2023-12-22]
29. Strayer DL, Cooper JM, Goethe RM, McCarty MM, Getty DJ, Biondi F. Assessing the visual and cognitive demands of in-vehicle information systems. Cogn Res Princ Implic 2019 Jun 21;4(1):18 [FREE Full text] [doi: 10.1186/s41235-019-0166-3] [Medline: 31227955]
30. Bérubé C, Schachner T, Keller R, Fleisch E, V Wangenheim F, Barata F, et al. Voice-based conversational agents for the prevention and management of chronic and mental health conditions: systematic literature review. J Med Internet Res 2021 Mar 29;23(3):e25933 [FREE Full text] [doi: 10.2196/25933] [Medline: 33658174]
31. Bérubé C, Fleisch E. Voice-based conversational agents for sensing and support: examples from academia and industry. In: Marsch LA, Jacobson NC, Kowatsch T, editors. Digital Therapeutics for Mental Health and Addiction: The State of the Science and Vision for the Future. Amsterdam, Netherlands: Elsevier Science; 2022:113-134.
32. Baki Kocaballi A, Laranjo L, Clark L, Kocielnik R, Moore RJ, Vera Liao Q, et al. Special issue on conversational agents for healthcare and wellbeing. ACM Trans Interact Intell Syst 2022 Jul 12;12(2):1-3 [FREE Full text] [doi: 10.1145/3532860]
33. Sezgin E, Militello LK, Huang Y, Lin S. A scoping review of patient-facing, behavioral health interventions with voice assistant technology targeting self-management and healthy lifestyle behaviors. Transl Behav Med 2020 Aug 07;10(3):606-628. [doi: 10.1093/tbm/ibz141] [Medline: 32766865]
34. Kahler CW, Lechner WJ, MacGlashan J, Wray TB, Littman ML. Initial progress toward development of a voice-based computer-delivered motivational intervention for heavy drinking college students: an experimental study. JMIR Ment Health 2017 Jun 28;4(2):e25 [FREE Full text] [doi: 10.2196/mental.7571] [Medline: 28659259]
35. Chung AE, Griffin AC, Selezneva D, Gotz D. Health and fitness apps for hands-free voice-activated assistants: content analysis. JMIR Mhealth Uhealth 2018 Sep 24;6(9):e174 [FREE Full text] [doi: 10.2196/mhealth.9705] [Medline: 30249581]
36. Kraus M, Wagner N, Minker W. Effects of proactive dialogue strategies on human-computer trust. In: Proceedings of the 28th ACM Conference on User Modeling, Adaptation and Personalization. 2020 Presented at: UMAP '20; July 14-17, 2020; Genoa, Italy URL: https://dl.acm.org/doi/abs/10.1145/3340631.3394840 [doi: 10.1145/3340631.3394840]

XSL•FO
RenderX

37. Dey AK, Abowd GD, Salber D. A conceptual framework and a toolkit for supporting the rapid prototyping of context-aware applications. Hum Comput Interact 2001;16(2-4) [FREE Full text] [doi: 10.1207/S15327051HCI16234_02]
38. Cha N, Kim A, Park CY, Kang S, Park M, Lee JG, et al. Hello there! Is now a good time to talk?: opportune moments for proactive interactions with smart speakers. Proc ACM Interact Mob Wearable Ubiquitous Technol 2020 Sep 04;4(3):1-28 [FREE Full text] [doi: 10.1145/3411810]
39. Zargham N, Reicherts L, Bonfert M, Voelkel ST, Schoening J, Malaka R, et al. Understanding circumstances for desirable proactive behaviour of voice assistants: the proactivity dilemma. In: Proceedings of the 4th Conference on Conversational User Interfaces. 2022 Presented at: CUI '22; July 26-28, 2022; Glasgow, UK URL: https://doi.org/10.1145/3543829.3543834 [doi: 10.1145/3543829.3543834]
40. Wu T, Martelaro N, Stent S, Ortiz J, Ju W. Learning when agents can talk to drivers using the INAGT dataset and multisensor fusion. Proc ACM Interact Mob Wearable Ubiquitous Technol 2021 Sep 14;5(3):1-28 [FREE Full text] [doi: 10.1145/3478125]
41. Lin CH, Shih HY, Sher PJ. Integrating technology readiness into technology acceptance: the TRAM model. Psychol Market 2007 May 24;24(7):641-657 [FREE Full text] [doi: 10.1002/mar.20177]
42. Dahlbäck N, Jönsson A, Ahrenberg L. Wizard of Oz studies: why and how. In: Proceedings of the 1st international conference on Intelligent user interfaces. 1993 Presented at: IUI '93; January 4-7, 1993; Orlando, FL URL: https://dl.acm.org/doi/10.1145/169891.169968 [doi: 10.1016/0950-7051(93)90017-n]
43. Minge M, Thuering M, Wagner I, Kuhr CV. The meCUE questionnaire: a modular tool for measuring user experience. In: Proceedings of the 7th Applied Human Factors and Ergonomics Society Conference 2016. 2016 Presented at: 7th Applied Human Factors and Ergonomics Society Conference 2016; July 27-31, 2016; Orlando, FL URL: https://www.researchgate.net/publication/305877075_The_meCUE_Questionnaire_A_Modular_Tool_for_Measuring_User_Experience [doi: 10.1007/978-3-319-41685-4_11]
44. Minge M, Thüring T. The MeCUE questionnaire (2.0): meeting five basic requirements for lean and standardized UX assessment. In: Proceedings of the Design, User Experience, and Usability: Theory and Practice. 2018 Presented at: Design, User Experience, and Usability: Theory and Practice; July 15-20, 2018; Las Vegas, NV URL: https://doi.org/10.1007/978-3-319-91797-9 [doi: 10.1007/978-3-319-91797-9_33]
45. Attig C, Wessel D, Franke T. Assessing personality differences in human-technology interaction: an overview of key self-report scales to predict successful interaction. In: Proceedings of the HCI International 2017. 2017 Presented at: HCI International 2017; July 9-14, 2017; Vancouver, BC. [doi: 10.1007/978-3-319-58750-9_3]
46. Attig C, Mach S, Wessel D, Franke T, Schmalfuß F, Krems JF. Technikaffinität als ressource für die Arbeit in industrie 4.0. In: Proceedings of the Innteract 2018 Conference. 2018 Jun 28 Presented at: Innteract 2018 Conference; 28/06/2018; Chemnitz, Germany. [doi: 10.14464/awic.v3i0.251]
47. Ren GJ, Szymanski MH, Arar R, Moore RJ. Studies in Conversational UX Design. Cham, Switzerland: Springer International Publishing; 2018.
48. Deci EL, Ryan RM. Self-determination theory: a macrotheory of human motivation, development, and health. Can Psychol 2008 Aug;49(3):182-185. [doi: 10.1037/a0012801]
49. Adell E. Acceptance of driver support systems. In: Proceedings of European Conference on Human Centred Design for Intelligent Transport Systems. 2010 Presented at: European Conference on Human Centred Design for Intelligent Transport Systems; April 29-30, 2010; Berlin, Germany.
50. Venkatesh V, Thong JY, Xu X. Consumer acceptance and use of information technology: extending the unified theory of acceptance and use of technology. MIS Q 2012;36(1):157-178 [FREE Full text] [doi: 10.2307/41410412]
51. Falkenström F, Hatcher RL, Skjulsvik T, Larsson MH, Holmqvist R. Development and validation of a 6-item working alliance questionnaire for repeated administrations during psychotherapy. Psychol Assess 2015 Mar;27(1):169-183 [FREE Full text] [doi: 10.1037/pas0000038] [Medline: 25346997]
52. Komiak SY, Benbasat I. The effects of personalization and familiarity on trust and adoption of recommendation agents. MIS Q 2006 Dec;30(4):941-960 [FREE Full text] [doi: 10.2307/25148760]
53. Parasuraman A, Colby CL. An updated and streamlined technology readiness index: TRI 2.0. J Serv Res 2014 Jun 22;18(1):59-74 [FREE Full text] [doi: 10.1177/1094670514539730]
54. Lehmann V, Zueger T, Maritsch M, Notter M, Bérubé C, Albrecht C, et al. 234-OR: noninvasive hypoglycemia detection during real car driving using in-vehicle data. Diabetes 2022 Jun 1;71(Supplement_1):234. [doi: 10.2337/db22-234-OR]
55. Bach KM, Jæger MG, Skov MB, Thomassen NG. Evaluating driver attention and driving behaviour: comparing controlled driving and simulated driving. In: Proceedings of the People and Computers XXII Culture, Creativity, Interaction. 2008 Presented at: People and Computers XXII Culture, Creativity, Interaction; September 1-8, 2008; Liverpool, UK. [doi: 10.14236/ewic/HCI2008.19]
56. Heise T, Koschinsky T, Heinemann L, Lodwig V, Glucose Monitoring Study Group. Hypoglycemia warning signal and glucose sensors: requirements and concepts. Diabetes Technol Ther 2003 Aug;5(4):563-571. [doi: 10.1089/152091503322250587] [Medline: 14511411]

XSL•FO
RenderX

57. Sun C, Malcolm JC, Wong B, Shorr R, Doyle MA. Improving glycemic control in adults and children with type 1 diabetes with the use of smartphone-based mobile applications: a systematic review. Can J Diabetes 2019 Feb;43(1):51-8.e3 [FREE Full text] [doi: 10.1016/j.jcjd.2018.03.010] [Medline: 30026048]
58. Feuerstein-Simon C, Bzdick S, Padmanabhuni A, Bains P, Roe C, Weinstock RS. Use of a smartphone application to reduce hypoglycemia in type 1 diabetes: a pilot study. J Diabetes Sci Technol 2018 Nov;12(6):1192-1199 [FREE Full text] [doi: 10.1177/1932296817749859] [Medline: 29291641]
59. Pagiatakis C, Rivest-Hénault D, Roy D, Thibault F, Jiang D. Intelligent interaction interface for medical emergencies: application to mobile hypoglycemia management. Smart Health 2020 Mar;15:100091. [doi: 10.1016/j.smhl.2019.100091]
60. The Nightscout Foundation. URL: https://www.nightscoutfoundation.org/ [accessed 2023-12-22]
61. Cox DJ, Singh H, Lorber D, Hermayer K. Diabetes and driving safety: science, ethics, legality and practice. Am J Med Sci 2013 Apr;345(4):263-265 [FREE Full text] [doi: 10.1097/maj.0b013e31828bf8d7]
62. Tobisch V. Why proactive voice AI will determine the Next In-Car UX. Cerence. 2020 Dec 9. URL: https://cerence.com/news-releases/news-release-details/why-proactive-voice-ai-will-determine-next-car-ux#:~:text=The%20concept%20of%20proactivity%20not,but%20have%20never%20reached%20them [accessed 2023-12-22]
63. Brito CF. Demonstrating experimenter and participant bias. In: Stowell JR, Addison WE, editors. Activities for Teaching Statistics and Research Methods: A Guide for Psychology Instructors. Washington, DC: American Psychological Association; 2017.
64. Maritsch M, Föll S, Lehmann V, Bérubé C, Kraus M, Feuerriegel S, et al. Towards wearable-based hypoglycemia detection and warning in diabetes. In: Proceedings of the Extended Abstracts of the 2020 CHI Conference on Human Factors in Computing Systems. 2020 Presented at: CHI EA '20; April 25-30, 2020; Honolulu, HI. [doi: 10.1145/3334480.3382808]
65. Heijmans M, Habets JG, Herff C, Aarts J, Stevens A, Kuijf ML, et al. Monitoring Parkinson's disease symptoms during daily life: a feasibility study. NPJ Parkinsons Dis 2019 Sep 30;5(1):21 [FREE Full text] [doi: 10.1038/s41531-019-0093-5] [Medline: 31583270]
66. Papathanail I, Vasiloglou MF, Stathopoulou T, Ghosh A, Baumann M, Faeh D, et al. A feasibility study to assess Mediterranean diet adherence using an AI-powered system. Sci Rep 2022 Oct 11;12(1):17008 [FREE Full text] [doi: 10.1038/s41598-022-21421-y] [Medline: 36220998]
67. Payne O, Smith V, Rolnik DL, Davies-Tuck M, Warty R, Seo D, et al. Virtual reality and its use in post-operative pain following laparoscopy: a feasibility study. Sci Rep 2022 Jul 30;12(1):13137 [FREE Full text] [doi: 10.1038/s41598-022-17183-2] [Medline: 35907935]
68. Kujala S, Roto V, Väänänen-Vainio-Mattila K, Karapanos E, Sinnelä A. UX Curve: a method for evaluating long-term user experience. Interact Comput 2011 Sep;23(5):473-483. [doi: 10.1016/j.intcom.2011.06.005]
69. Dell N, Vaidyanathan V, Medhi I, Cutrell E, Thies W. "Yours is better!": participant response bias in HCI. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. 2012 Presented at: CHI '12; May 5-10, 2012; Austin, Texas URL: https://doi.org/10.1145/2207676.2208589 [doi: 10.1145/2207676.2208589]

Abbreviations

ATI: Affinity for Technology Interaction
AUT: Acceptance and Use of Technology
CGM: continuous glucose monitoring
FGM: flash glucose monitoring
meCUE: modular evaluation of key Components of User Experience
RQ: research question
SAI: Session Alliance Inventory
T1DM: type 1 diabetes mellitus
TRI: Technology Readiness Index

Edited by A Kushniruk; submitted 15.12.22; peer-reviewed by M Chatzimina; comments to author 31.01.23; revised version received 06.02.23; accepted 04.12.23; published 09.01.24.

Please cite as:
Bérubé C, Lehmann VF, Maritsch M, Kraus M, Feuerriegel S, Wortmann F, Züger T, Stettler C, Fleisch E, Kocaballi AB, Kowatsch T
Effectiveness and User Perception of an In-Vehicle Voice Warning for Hypoglycemia: Development and Feasibility Trial
JMIR Hum Factors 2024;11:e42823
URL: https://humanfactors.jmir.org/2024/1/e42823
doi:10.2196/42823
PMID:38194257

XSL•FO
RenderX

©Caterina Bérubé, Vera Franziska Lehmann, Martin Maritsch, Mathias Kraus, Stefan Feuerriegel, Felix Wortmann, Thomas Züger, Christoph Stettler, Elgar Fleisch, A Baki Kocaballi, Tobias Kowatsch. Originally published in JMIR Human Factors (https://humanfactors.jmir.org), 09.01.2024. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on https://humanfactors.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

A Closed-Loop Falls Monitoring and Prevention App for Multiple Sclerosis Clinical Practice: Human-Centered Design of the Multiple Sclerosis Falls InsightTrack

Valerie J Block[1,2], PT, DPTSc; Kanishka Koshal[1], BDS, MSc; Jaeleene Wijangco[1], BS; Nicolette Miller[1], BS; Narender Sara[1], MBA, MS; Kyra Henderson[1], BA; Jennifer Reihm[1], MPA; Arpita Gopal[2], DPT, PhD; Sonam D Mohan[1], MD, PhD; Jeffrey M Gelfand[1], MAS, MD; Chu-Yueh Guo[1], MD; Lauren Oommen[1], DO; Alyssa Nylander[1], MD, PhD; James A Rowson[1], PhD; Ethan Brown[1], MD; Stephen Sanders[1], PhD; Katherine Rankin[1], PhD; Courtney R Lyles[3,4], PhD; Ida Sim[5], MD, PhD; Riley Bove[1], MMSc, MD

[1]Department of Neurology, University of California San Francisco Weill Institute, University of California San Francisco, San Francisco, CA, United States

[2]Department of Physical Therapy and Rehabilitation Science, University of California San Francisco, San Francisco, CA, United States

[3]University of California San Francisco Division of General Internal Medicine, Zuckerberg San Francisco General Hospital, San Francisco, CA, United States

[4]Center for Vulnerable Populations, University of California San Francisco, San Francisco, CA, United States

[5]Department of Medicine, University of California San Francisco, San Francisco, CA, United States

Corresponding Author:
Riley Bove, MMSc, MD
Department of Neurology
University of California San Francisco Weill Institute
University of California San Francisco
Box 3126 1651 4th St
Room 612A
San Francisco, CA, 94143
United States
Phone: 1 (415) 353 2069
Email: riley.bove@ucsf.edu

Abstract

Background: Falls are common in people with multiple sclerosis (MS), causing injuries, fear of falling, and loss of independence. Although targeted interventions (physical therapy) can help, patients underreport and clinicians undertreat this issue. Patient-generated data, combined with clinical data, can support the prediction of falls and lead to timely intervention (including referral to specialized physical therapy). To be actionable, such data must be efficiently delivered to clinicians, with care customized to the patient's specific context.

Objective: This study aims to describe the iterative process of the design and development of Multiple Sclerosis Falls InsightTrack (MS-FIT), identifying the clinical and technological features of this closed-loop app designed to support streamlined falls reporting, timely falls evaluation, and comprehensive and sustained falls prevention efforts.

Methods: Stakeholders were engaged in a *double diamond* process of human-centered design to ensure that technological features aligned with users' needs. Patient and clinician interviews were designed to elicit insight around ability blockers and boosters using the capability, opportunity, motivation, and behavior (COM-B) framework to facilitate subsequent mapping to the Behavior Change Wheel. To support generalizability, patients and experts from other clinical conditions associated with falls (geriatrics, orthopedics, and Parkinson disease) were also engaged. Designs were iterated based on each round of feedback, and final mock-ups were tested during routine clinical visits.

Results: A sample of 30 patients and 14 clinicians provided at least 1 round of feedback. To support falls reporting, patients favored a simple biweekly survey built using REDCap (Research Electronic Data Capture; Vanderbilt University) to support *bring-your-own-device* accessibility—with optional additional context (the severity and location of falls). To support the evaluation and prevention of falls, clinicians favored a clinical dashboard featuring several key visualization widgets: a longitudinal falls

XSL•FO
RenderX

display coded by the time of data capture, severity, and context; a comprehensive, multidisciplinary, and evidence-based checklist of actions intended to evaluate and prevent falls; and MS resources local to a patient's community. In-basket messaging alerts clinicians of severe falls. The tool scored highly for usability, likability, usefulness, and perceived effectiveness (based on the Health IT Usability Evaluation Model scoring).

Conclusions: To our knowledge, this is the first falls app designed using human-centered design to prioritize behavior change and, while being accessible at home for patients, to deliver actionable data to clinicians at the point of care. MS-FIT streamlines data delivery to clinicians via an electronic health record–embedded window, aligning with the *5 rights* approach. Leveraging MS-FIT for data processing and algorithms minimizes clinician load while boosting care quality. Our innovation seamlessly integrates real-world patient-generated data as well as clinical and community-level factors, empowering self-care and addressing the impact of falls in people with MS. Preliminary findings indicate wider relevance, extending to other neurological conditions associated with falls and their consequences.

(JMIR Hum Factors 2024;11:e49331) doi:10.2196/49331

KEYWORDS

digital health; mobile tools; falls; prevention; behavioral medicine; implementation science; closed-loop monitoring; multiple sclerosis; mobile phone

Introduction

Background

Falls are common in patients with multiple sclerosis (MS), occurring in 50% to 70% of published cohorts, a rate similar to that of older adults [1]. Falls often lead to injury, result in significant health care costs [2-5], and increase *thefear of falling* [6,7]; furthermore, they lead to a decline in physical activity and participation in daily life as well as cause loss of independence [8,9]. Targeted interventions such as physical therapy (PT) can reduce falls and the fear of falling [10-12], but patients often underreport and clinicians undertreat this issue. Indeed, fewer than half of the people with MS who report falls receive falls prevention information from their clinician [13], and there is a lack of self-management apps to engage and empower people with MS about falls prevention [14-16].

To address this gap, multimodal closed-loop tools hold promise. Closed-loop tools can use real-time feedback and patient-generated data (PGD; such as those already validated in MS [17-22]) to continuously monitor and adjust interventions to improve outcomes. Such an approach has been used in biological functions and symptoms, such as insulin delivery or depression [23-25]. Unfortunately, in MS, apps on the commercial market exist outside of the health system, that is, away from the point of care. To close these gaps in care, a tool should close the loop of information flow from the patient to the appropriate clinician (depending on the diagnosis and symptoms being treated, ie, neurologist) at the point of care and back to the patient to support patient-centered care. Furthermore, the tool must address the behavioral barriers to change to promote the behaviors (eg, reporting, screening, treatment recommendations, and follow-up with timely refills or referral scheduling) likely to lead to falls prevention. From previous work, real-time PGD such as prospective near-falls reports, patient-reported outcomes, and changes in step count captured by wearable sensors all provide useful input for the closed-loop models [26,27]. The integration of these in a multimodal tool would enhance falls prediction accuracy and could act as an early warning system for timely PT referrals, reducing falls risk and related injuries [28,29]. However, challenges lie in delivering PGD to the point of care, granting access for prompt intervention, and active self-management. To be actionable, these PGD, generated from remote devices or patient-reported outcomes, must be delivered according to the *5 rights* [30]: the right information, to the right person, in the right format, through the right channel, at the right time in the workflow. This is a hurdle that health systems have for the most part not yet overcome, and PGD are not typically integrated into care systems.

To address these challenges, we developed Multiple Sclerosis Falls InsightTrack (MS-FIT), a closed-loop falls monitoring and prevention app. MS-FIT enables seamless information exchange between patients and clinicians, driven by stakeholder input and human-centered design (HCD) principles [31,32]. It empowers individuals with MS to *track* falls, enhances clinician decision-making by providing real-world *insights*, and fills a crucial gap in self-management for falls monitoring and prevention.

Objectives

This paper describes the iterative process of the design and development of MS-FIT. MS-FIT is designed to integrate various data types to personalize falls risk assessments and interventions for individuals with MS. To achieve this, a planned process of engagement of patients and clinicians (ie, neurologists) was performed to ensure that MS-FIT aligns with user needs, whereas usability evaluations validated its potential impact on falls prevention. Subsequently, we will test the feasibility of implementation and effectiveness of MS-FIT in a larger clinical trial.

Methods

Study Setting

The primary clinical setting is the University of California San Francisco (UCSF) Multiple Sclerosis and Neuroinflammation Center, which provides specialized care to >6000 adults with MS annually. Clinician stakeholders were approached via email or in person and invited to participate in the study. Patients who had given permission to be contacted for research participation

XSL•FO
RenderX

or who had sustained falls in the past year were invited via secure email to participate as stakeholders.

Ethical Considerations

The University of California San Francisco Institutional Review Board approved all study activities (22-36680). Informed consent forms and Health Insurance Portability and Accountability Act documents were signed by each study participant (patients, clinicians, and other interviewees). Patients received US $50 (1-time compensation) for their participation in the study.

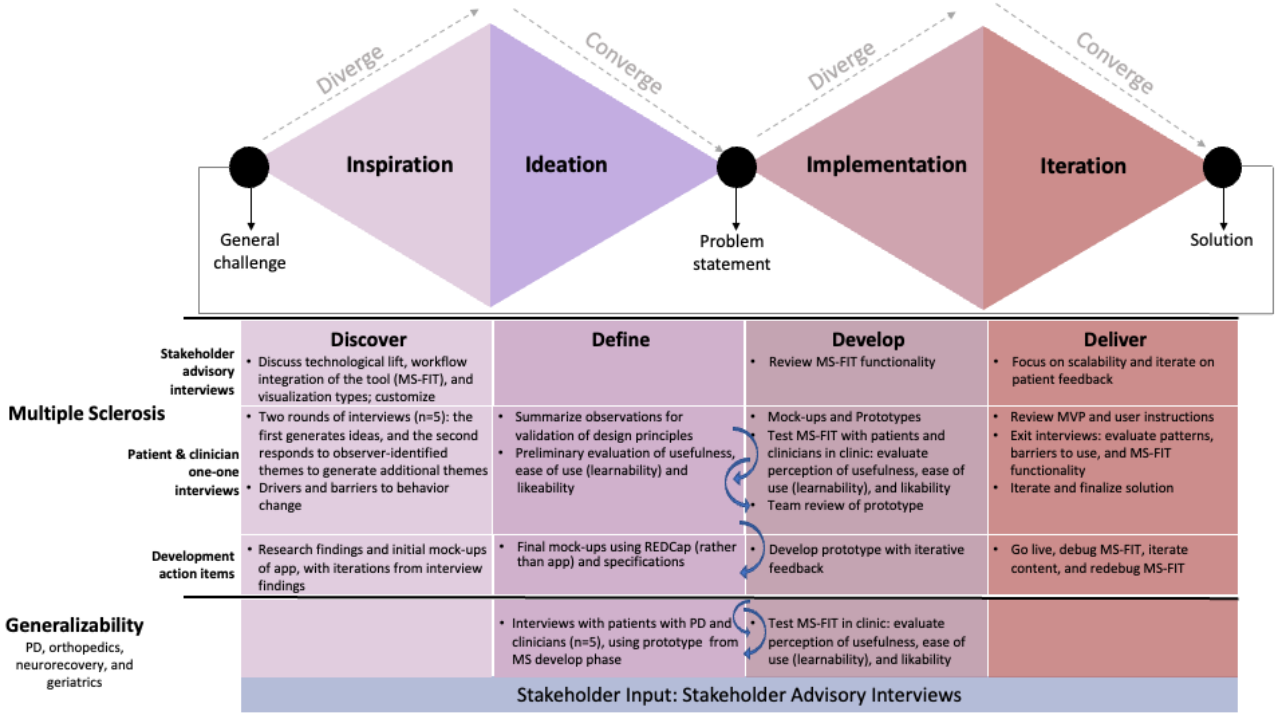
Study Design

The overarching approach was grounded in the principles and phases of HCD [31]. This process focuses on the usability and needs of those whom the tool is meant to serve, in this case, patients and clinicians. The development protocols included (1) thorough engagement from a comprehensive range of stakeholders, (2) models based on HCD approaches to ensure alignment with the needs of the intended users (patients and clinicians), (3) an evaluation of the tool's usability using an established framework: the Health IT Usability Evaluation Model (Health-ITUEM) [33], and (4) plans to support the generalizability and scalability of the tool to other clinical settings associated with falls.

HCD involves a series of steps, articulated initially in the context of design [34] and expanded to health care [35]: inspire (empathize with all stakeholders), ideate (define the problem and conceptualize in an open-minded manner), implement (prototype solutions and test), and iterate. Figure 1 illustrates these phases in a *modified double diamond* approach as they were undertaken in the current project, depicting the iterative broadening and narrowing of content and layout throughout the phases [36]. Figure 2 shows the trajectory of MS-FIT and the assimilation of insights obtained from user interviews (involving patients and clinicians) throughout the phases of discover, define, develop (iterative), and deliver.

The initial prototype (prediscover) was developed based on feedback from extensive HCD of the BRIDGE point-of-care clinical dashboard (refer to the Technological Building Blocks subsection) summarized elsewhere [37,38], where both patients and clinicians expressed a desire for the integrations of features and episodes of falls to be incorporated into the design. The study team initially identified key elements for MS-FIT through a combination of clinical expertise and literature review [39,40] (Figure 3). These elements were then amalgamated into mock app screens using PowerPoint (Microsoft Corp) for the first round of patient interviews. Figure 3 illustrates the inaugural prototype, which was informed by valuable insights from observational [41] and interventional [39] studies that used PGD to monitor walking and falls in individuals with MS. In addition, the prototype draws inspiration from clinician-facing [42] and patient-facing [43] apps designed using HCD principles to promote shared decision-making and evidence-based practice in MS.

Figure 1. Modified double diamond approach: phases of development and stakeholder engagement. The double diamond depicts the human-centered design principles and framework, with iterations through the discover, define, develop, and delivery phases. The timeline and workflow of the human-centered design phases depict corresponding interviews and products. The curved arrows between "Define" and "Develop" indicate an iterative process between these 2 phases. MS: multiple sclerosis; MS-FIT: Multiple Sclerosis Falls InsightTrack; MVP: minimum viable product; PD: Parkinson disease; REDCap: Research Electronic Data Capture.



XSL•FO
RenderX

Figure 2. The trajectory of Multiple Sclerosis Falls InsightTrack (MS-FIT) though the phases of development and stakeholder engagement. The final tool components include a patient survey (MS-FIT patient survey) and a clinical dashboard (MS-FIT BRIDGE). The trajectory integrates feedback from user (patient and clinician) interviews through the phases of discover, define, develop (iterative), and deliver. The version numbers indicate a revised version of the patient- or clinician-facing prototype. "Other patients" refers to patients with Parkinson disease as well as orthopedics, neurorecovery, and geriatrics populations. "Full test" refers to the prototype testing in the contextual environment. MS: multiple sclerosis.

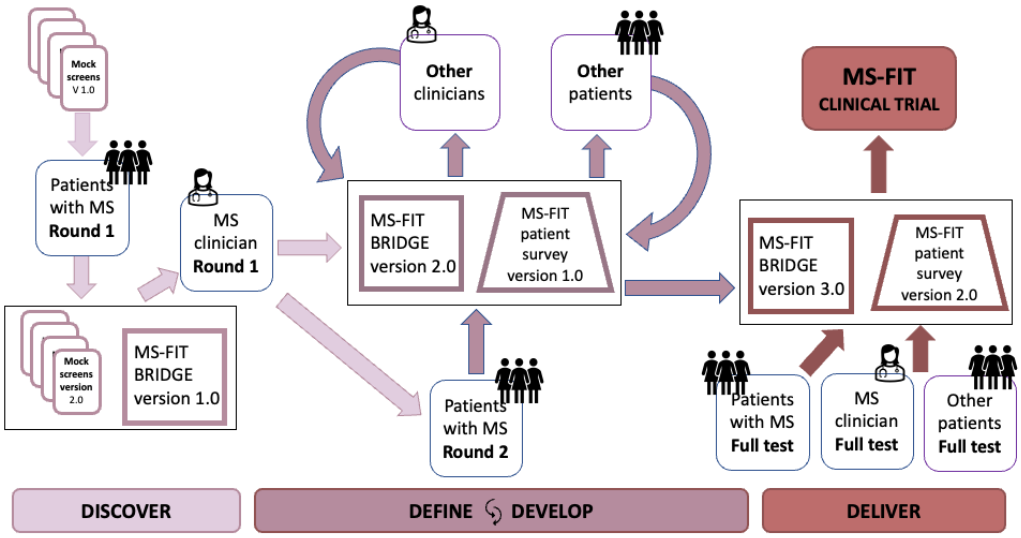
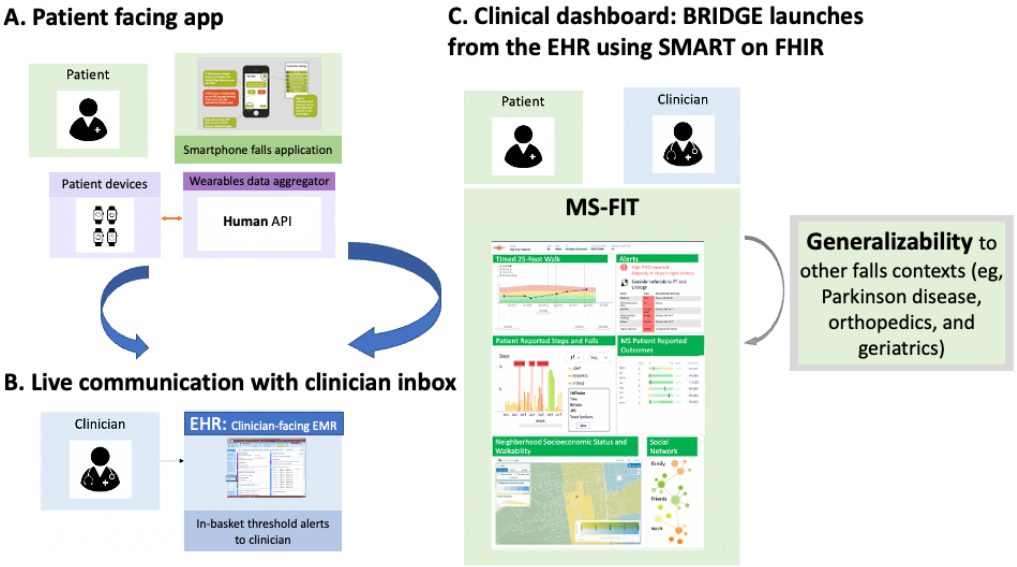


Figure 3. Initial proposal for Multiple Sclerosis Falls InsightTrack (MS-FIT), which involved designing a closed-loop integrated MS-FIT personal health library. MS-FIT is designed to enable patients to track their falls in the context of their lived experience, report them to their care team, and gain insight into multimodal contributors to falls, falls' impact on daily life participation, and interventions likely to prevent falls. Clinicians, by using BRIDGE, can gain insight into which patients are falling between clinical encounters and how best to personalize risk reduction interventions for the individual patient. This prototype was generated from a number of insights from observational and interventional studies that used patient-generated data to monitor walking and falls in people with multiple sclerosis (MS) and from clinician-facing [42] and patient-facing [43] apps designed using human-centered design to facilitate shared decision-making and evidence-based practice in MS. (A) Patient facing app; (B) Live communication with clinician inbox; (C) Clinical dashboard: BRIDGE launches from the EHR using SMART or FHIR. API: application programming interface; EHR: electronic health record; EMR: electronic medical record; FHIR: Fast Healthcare Interoperability Resources; MS-FIT: Multiple Sclerosis Falls InsightTrack; SMART: Substitutable Medical Apps and Reusable Technologies.



Framework for Tool Evaluation

The Health-ITUEM framework appraises both subjective and objective outcomes that inform a tool's usability [44]. In the design phases described herein, the subjective outcomes (satisfaction measured by the perceived ease of use and perceived usefulness) were primarily evaluated. Furthermore, the 4 key variables proposed by Mathews et al [45] to determine both (1) whether the tool (MS-FIT) reflects HCD principles and (2) whether it is likely to engage patients were applied. These four domains encompass (1) usefulness, (2) ease of use or learnability, (3) likability, and (4) effectiveness. These frameworks were used to categorize critical data and visualization elements, as well as the technological and clinical workflow aspects of MS-FIT [46].

XSL•FO
RenderX

Technological Building Blocks

The architecture of the tool was built leveraging existing tools, primarily BRIDGE and REDCap (Research Electronic Data Capture; Vanderbilt University).

BRIDGE

The BRIDGE precision medicine platform at UCSF is an application programming interface (API) that assembles clinical and research data from a variety of sources into a dashboard customized for a given clinical context, displaying a series of digestible, actionable visualizations [38]. BRIDGE is integrated with the Epic electronic health record (EHR; Epic Systems Corporation), launches from Epic using Substitutable Medical Apps and Reusable Technologies (SMART) on Fast Healthcare Interoperability Resources (FHIR; a standard approach for building reusable and extendable EHR-integrated apps), and is integrated with Epic FHIR APIs and other data integrations. The back-end of BRIDGE is built using Python, the flask framework, and PostgreSQL to store configuration data. Although individual-level data will populate the tool, cohort-level data can become the reference cohort against which an individual's data can be contextualized. BRIDGE pulls data not only from the EHR but also from a range of custom research databases as well as other APIs, such as REDCap [38]. BRIDGE was developed based on extensive HCD processes both within the field of MS [42,43] and beyond [37]. The data visualizations can be developed using HTML, cascading style sheets (CSS), JavaScript, Data-Driven Documents–JavaScript, and other front-end libraries. Each front-end visualization is modular, allowing for asynchronous loading, and is a parameterized JavaScript component, allowing us to extend the code to additional platforms and data sources. Data formatting standards are also applied to make all visualizations and data inputs modular. All API calls are made in real time; BRIDGE does not store patient data, but there is an option to write back to the EHR by pasting the visuals into a clinical note. Furthermore, the development team follows universal design principles, influenced by the Agency for Healthcare Research and Quality Toolkit for Designing Consumer Health IT [47].

REDCap Tool

REDCap [38] includes editable or annotatable functions to enable patients to keep track of, and annotate, their PGD. Design choices reflect digital health literacy principles and feedback provided from diverse patients. Together, these enhancements make the data understandable and actionable.

Investigator Team

The core team included an MS neurologist with HCD expertise (RB), software engineers (NM and NS), a health literacy and patient engagement expert (JR), and an MS physical therapist with remote ambulatory and falls monitoring expertise (VJB). Additional key scientific input was provided by a digital health cloud infrastructure expert (IS), an implementation science expert (CL), a health disparities and population health expert (CL), and an expert in large-scale mobile health (IS). Patient stakeholders included National Multiple Sclerosis Society advocates (LG) and patients (3 core stakeholders). Research team members included a program manager (KK) and clinical coordinators (JW and KH). Before starting the project, this team met to determine the phases of research and design an initial mock-up of the tool that could be used during the discover phase. Volunteer consultants included a software engineer (JR) and user interface or user experience experts.

Phases of Design

Phase 1: Discover

Stakeholder Advisory Team

An initial stakeholder meeting took place, during which the goals and phases of the project were outlined. Next, the core team met biweekly as a group or as subgroups to discuss an agenda that included the development of patient and clinician interview guides, interview coding schemes and thematic analysis, the practical aspects of the technological lift, the workflow integration of MS-FIT, and the visualization types and customizations. The iterations of mock-ups were revised based on patient and clinician interview feedback.

Interviews

One-on-one interviews were conducted by the health literacy and patient engagement expert (hereinafter referred to as the interviewer) with patients (round 1) and clinicians. Because of ongoing COVID-19 restrictions on in-person engagements, interviews were conducted via the UCSF Zoom video platform (Zoom Video Communications, Inc) using interview guides developed for each audience to elucidate how a tool might be designed to promote behavior change around falls ascertainment, reporting, and prevention. All questions were administered verbally, and interviews lasted between 45 and 60 minutes. With participant consent, interviews were simultaneously recorded and transcribed using Zoom's video transcription feature.

Interview guides included qualitative and quantitative components. Open-ended questions probed around the domains of the capability, opportunity, motivation, and behavior (COM-B) framework to facilitate subsequent mapping to the Behavioral Change Wheel (BCW) proposed by Michie et al [48]. Quantitative questions with Likert-style responses (ranging from 1=lowest to 5=highest) were administered verbally throughout each interview to assess specific aspects of patient and clinician experience related to capability, opportunity, and motivation, as well as the perceived usefulness of mock screen views and workflows. Participants were asked to comment on their Likert-style responses.

Patient interviews were semistructured around 2 key thematic topics: patient experience with (1) falls and activity, including ability to be active, knowledge, communication with care team, experience, feelings, and expectations; and (2) use of technology, including smartphone, tracking devices, apps, and communication with care team. To complement qualitative insights, patients were asked to use a Likert scale to rate the perceived usefulness of each of 3 app screen views featuring different design elements.

Semistructured interviews with clinicians started with a review of the activity blockers and boosters identified during the discover phase interviews with patients. With this insight,

XSL•FO
RenderX

clinicians were asked a series of open-ended qualitative questions to elicit their perspectives on whether a falls reporting tool might promote sustainable falls prevention, as well as gather feedback on the initial closed-loop design (Figure 3) intended to support falls treatment and clinical decision support. To assess each design feature, clinicians were asked to rate perceived usefulness on a Likert scale.

Analysis

After all interviews were concluded for each audience, the interviewer reviewed each transcript and used inductive coding to develop a coding scheme on the basis of responses to the open-ended questions [46]. Frequently occurring topics were assigned a unique thematic category, and less frequent topics were coded *other*. Categories were defined by the interviewer, and quotations from the transcript were used to illustrate the type of text coded into the category. Although the interviewer was the sole coder, the stakeholder advisory team provided ongoing consultation on the coding scheme and how to code less frequently occurring responses.

The interviewer transferred Likert-style response data to a spreadsheet to calculate means and SDs for each question. To analyze questions designed to map to the COM-B framework, the interviewer created a data grid where the rows were COM-B categories with subthemes of ability blocker and booster types, and the columns were evidence (quotes) of specific blockers or boosters [49]. Evidence of blockers or boosters that spanned >1 category were placed in all relevant categories to ensure that they would be represented when considering BCW-guided interventions.

After developing the initial COM-B data grid, the interviewer, in consultation with the stakeholder advisory group, expanded the grid to include (1) BCW intervention functions to help users overcome barriers to performing target behaviors and (2) potential intervention solution features designed to be effective for each corresponding blocker category. Intervention solution features were subsequently added to the design road map for immediate or future implementation.

Phases 2 and 3: Define and Develop (Iterative)

Stakeholder Advisory Team

In these phases, the team reviewed qualitative and quantitative findings from additional patient (2 rounds) and clinician (1 round) interviews and used this feedback to further refine MS-FIT tool functionality, including design and technological features. Changes were prioritized according to the strength of feedback (occurrence of themes and usability scores) and technical feasibility.

Interviews

The define and develop phases encompassed a second round of patient interviews, followed by 2 rounds of interviews with clinicians and patients designed to assess MS-FIT generalizability to other high-risk clinical contexts. The same process was followed as that described in phase 1 (discover). One-on-one interviews were conducted by the interviewer via the UCSF Zoom video platform using interview guides. All questions were administered verbally, and interviews lasted between 45 and 60 minutes. With participant consent, interviews were simultaneously recorded and transcribed using Zoom's video transcription feature.

Patient Interviews (Round 2)

Interview guides included qualitative and quantitative components. In an effort to validate the patient experience findings from round 1 interviews, patients interviewed during round 2 were similarly asked to share qualitative feedback around personal experiences with falls, falls and near-falls reporting, perceived benefits and concerns around using a falls tracking app, and thoughts on what supports would be helpful between appointments. Quantitative questions with Likert-style responses (ranging from 1=lowest to 5=highest) were used to rate 9 mock screens for usefulness, understandability, and importance for each view. Mock screens had been iterated after the discover phase; therefore, patient feedback during this second round further validated and helped refine the designs.

Generalizability to Other High-Risk Clinical Contexts

To ensure that the technological build was not *overdesigned* for MS and to support the scalability of the tool to other clinical settings, interviews were expanded to intended users in other clinical specialties associated with falls, including geriatrics, orthopedics, neurorecovery (after stroke or traumatic brain injury), and Parkinson disease (PD). Clinicians from each discipline and patients with PD were interviewed. Interview protocols used during the discover phase were adapted to reference specific disciplines and diseases, whereas the questions (qualitative and quantitative) remained the same to yield a parallel assessment of each audience's experiences, preferences, capabilities, opportunities, and motivations.

Analysis

Qualitative and quantitative interview analysis used the same inductive coding and calculation techniques, respectively, used during the discover phase. The results were analyzed by the interviewer, with ongoing thematic consultation with the stakeholder advisory team, and used to inform and prioritize design and content iterations.

Phase 4: Deliver

Stakeholder Advisory Team

The core team met with stakeholders on an ad hoc small-group basis during this phase to plan observation and tool-scoring protocols, specifically to identify a subset of questions from the Health IT Usability Evaluation Scale (Health-ITUES) derived from the Health-ITUEM to assess the 2 *subjective* components of usability—usefulness and ease of use [33]—as well as the Patient Education Materials Assessment Tool for Audiovisual Materials to assess understandability and actionability [50]. As recommended for digital tool validation [45], a single survey question—Net Promoter Score (NPS)—was asked regarding the likelihood that users (patients and clinicians) would recommend the MS-FIT to colleagues or friends. Additional conversations focused on the scalability of the tool, as well as the qualitative and quantitative feedback received.

XSL•FO
RenderX

Observations and Scoring

Observations and scoring for the patient-facing falls assessment survey took place with 2 audiences: people with MS and people with PD. Patients scheduled for a routine upcoming in-person clinical visit with their neurologist were contacted and invited to participate in testing and evaluating the tool. After providing informed written consent, and while being observed by the interviewer, participants were asked to engage with the MS-FIT minimum viable product consisting of the falls assessment survey and accompanying patient instructions while being observed by the interviewer. Patients were specifically asked to complete the falls assessment survey by entering up to 5 falls (real or hypothetical) that had occurred in the prior 2 weeks and responding to on-screen prompts to provide context about each reported fall. Patients could ask questions of the interviewer, if needed. After survey submission, each patient was asked to complete an 18-item survey about their experience to assess usability, usefulness and ease of use, likability, understandability, actionability, and NPS. Patients were subsequently asked if they had any feedback about their experience. Feedback was documented in field notes captured contemporaneously.

Clinicians seeing people with MS and those with PD who had just been observed entering data in the falls assessment survey were asked to launch the MS-FIT BRIDGE app in real-time clinical encounters with these patients to review the falls and contextual data the patient had entered and to engage with the various widgets designed to help evaluate and address reported falls. The interviewer met with the clinicians immediately after the encounters to conduct in-person exit interviews and administer a 9-item survey to assess usability, usefulness and ease of use, likability, understandability, actionability, and NPS. Clinicians were subsequently asked whether they had any feedback about their experience, including any barriers to use and functionality challenges. Feedback was documented in contemporaneous field notes.

Analysis

Qualitative feedback, although limited, was analyzed by the interviewer using the same inductive coding technique used during the previous 2 interview phases. Quantitative questions with Likert-style responses (ranging from 1=lowest to 5=highest) were used to score likability, usability, usefulness, and ease of use. Understandability and actionability were assessed using a binary *agree* or *disagree* scale. Another member of the research team entered quantitative responses into REDCap, which was used to calculate means and SDs for all Likert-style responses and total binary responses. NPS responses (0-10 scale) were calculated by the interviewer by subtracting the percentage who were detractors (those who scored 0 to 6) from the percentage who were promoters (those who scored 9 or 10). An NPS >0 was considered good, >20 was considered favorable, and >50 was considered excellent.

Development Action Items

Once the tool was live, the developer was able to debug MS-FIT; iterate based on patient, clinician, and stakeholder feedback; and redebug as needed.

Results

Overview

Demographic information about each interview panel is shown in Multimedia Appendix 1 [42]. Altogether, 30 patients of diverse ages, disability levels, and technological literacy as well as 14 clinicians provided at least 1 round of feedback. The level of involvement from the users ranged from testers to informants [32]. Feedback from both rounds of interviews with people with MS, MS clinician comments, and feedback from other high-risk clinical context patient and clinician interviews were integrated into the final MS-FIT design. Iterative interview feedback was categorized into activity blockers (what keeps people from performing a behavior) and boosters (what is already working well that we can build on) in the COM-B model. Examples of how interview feedback findings fit into the COM-B and BCW, along with intervention function solution features integrated into MS-FIT, are shown in Figure 4. Details are provided in Table 1. Further discussions with clinicians and patients in other high-risk clinical contexts confirmed the findings from the MS context. Across these specialties, the main barriers to falls prevention efforts included access to specialized PT (availability and physical ability to access it), insurance coverage, ability to adapt the home to improve safety, the adequate use of assistive devices, and COVID-19–related restrictions to community exercise areas.

The overview of findings from interviews with clinicians and participants with MS, highlighting areas that block or boost patient and clinician behavior change with regard to falls and falls prevention, are shown in Multimedia Appendix 2.

XSL•FO
RenderX

Figure 4. Example of mapping blockers and boosters relating to falls prevention (findings from interviews with patients with multiple sclerosis [MS]) to the Behavioral Change Wheel [49] and associated behaviorally informed intervention solution features. Examples for each of the sections of the capability, opportunity, motivation, and behavior (COM-B) model are highlighted, showing how these integrate into the Behavioral Change Wheel. The examples provided relate to patients' reported goals, blockers (features that block falls prevention behavior), and boosters (features that boost behaviors related to falls prevention).

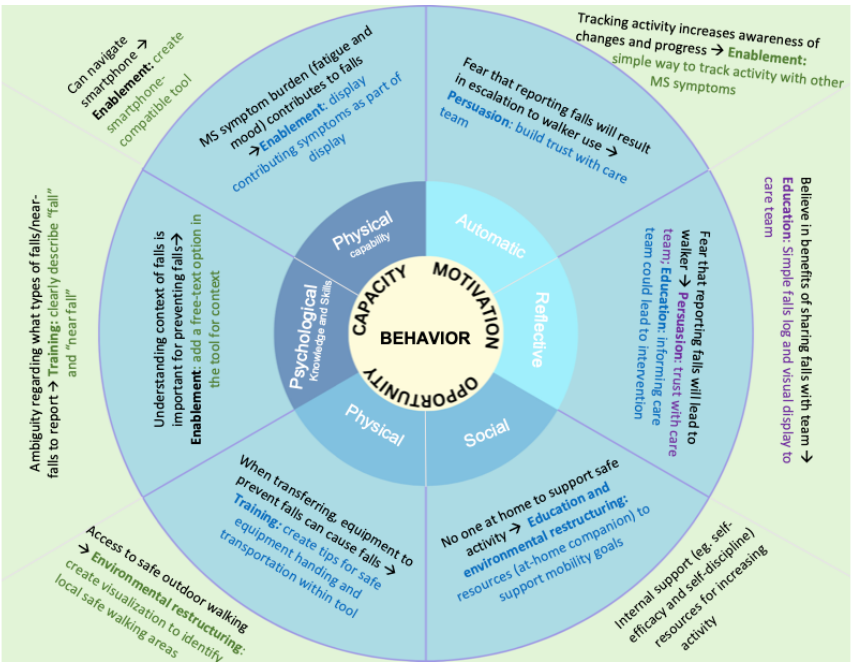


Table 1. Scoring of the final University of California San Francisco Multiple Sclerosis Falls InsightTrack app (REDCap [Research Electronic Data Capture]) by patients with multiple sclerosis: usability, ease of use, and likability (n=10).

Health-ITUES[a]–based questions for usability, ease of use, and likability	Score, mean (SD)	Score <4 out of 5, n (%)
"It is useful to report if I've had any falls or near falls every 2 weeks"	4.80 (0.42)	0 (0)
"It is useful to have my survey answers sent to my care team"	4.90 (0.32)	0 (0)
"The survey asks about important topics"	4.70 (0.48)	0 (0)
"I am comfortable with my ability to complete the survey"	4.80 (0.42)	0 (0)
"I find the survey easy to use"	4.80 (0.42)	0 (0)
"I can easily remember how to access the survey through my email"	4.60 (0.70)	1 (10)
"I like the survey"	4.80 (0.42)	0 (0)

[a]Health-ITUES: Health IT Usability Evaluation Scale (scores range from 1=strongly disagree to 5=strongly agree).

Tool Components

Thematic saturation was reached after 5 patient interviews (round 1), and we incorporated these insights into prototypes for an additional 5 patient interviews (round 2), which were then iteratively reviewed.

UCSF Support Self-Monitoring: A Patient-Facing Tool to Track Falls and Self-Monitor

Tool Architecture

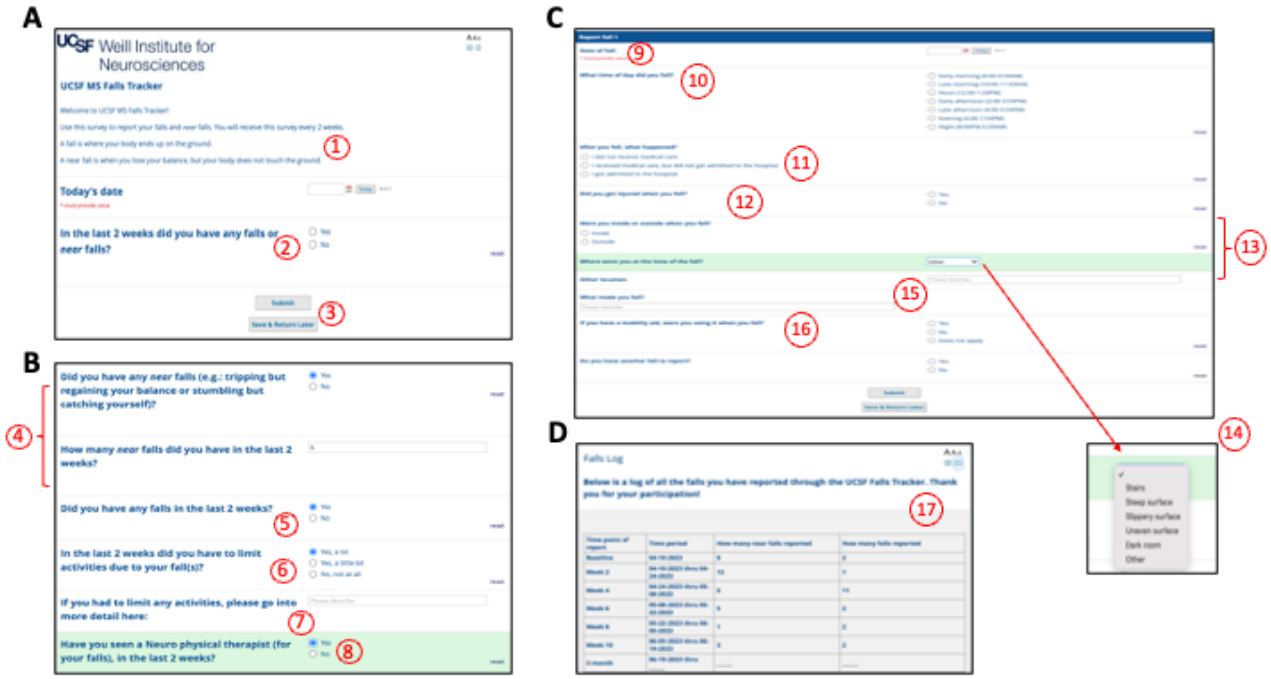
One key and consistent theme emerging from patient interviews was a preference for a simpler design for the patient-facing tool than had been initially conceived. Combined with a goal of maintaining confidentiality and keeping personal information within our university firewall, the study team opted for a tailored REDCap app rather than a custom new app.

Tool Components

Key features informed by patient and clinician feedback are detailed in Figure 5. Key features mapping to the COM-B framework (Multimedia Appendix 2 and Figure 4) are denoted by a red number and described in Textbox 1.

XSL•FO
RenderX

Figure 5. University of California San Francisco (UCSF) Multiple Sclerosis (MS) Falls Tracker: a patient-facing tool to track falls and support self-monitoring. This is the "MS-FIT [Multiple Sclerosis Falls InsightTrack] patient survey V2.0," sent via email to patients with a REDCap (Research Electronic Data Capture) survey link. Key features mapping to user-generated perspectives and feedback and to the capability, opportunity, motivation, and behavior (COM-B) framework are denoted by a red number and described in Textbox 1.



XSL•FO
RenderX

Textbox 1. University of California San Francisco Multiple Sclerosis Falls InsightTrack: key features. The numbers correspond to the red numbers in Figure 5, which denote key features mapping to the capability, opportunity, motivation, and behavior (COM-B) framework.

Concise and precise falls screening

1. Clear definitions were preferred to distinguish between a fall and a near fall to support the reporting of meaningful data.
2. An easy-to-use and simple 1-question tool that could be completed frequently (every 2 wk) was preferred to relying on "flawed memory" to report falls during sporadic clinic visits: if "No," then the survey ends at this point; if "Yes," then branching logic continues.
3. The ability to easily report each fall or near fall separately was preferred. The ability to edit (return later) was important for reducing burden.
4. Simple reporting for near falls (yes or no and overall number) was preferred, given the large volume of near falls experienced by some patients and the potential burden and time commitment of providing details.
5. The 2-wk epoch between reporting was determined feasible (balance between memory and overburdening).
6. The ability to report activity limitations was preferred because these pertain to primary goals with regard to the "ability to continue independence for activities of daily living" and to "stay active."
7. Because of the heterogeneity in answers, a free-text option would allow patients to add further details regarding activity limitations.
8. Indicating whether the patient has seen a neurorehabilitation specialist could help clinicians triage the continued plan of care.

Detailed context of falls (optional)

1. Recording the date of the fall using a simple button allows the tool to display each fall into the longitudinal representation (refer to Textbox 2; Figure 6).
2. The time of falls can also inform falls context (eg, in the dark or when fatigued). The 24-h day was divided into time blocks for clarity and to reduce recall error of exact time.
3. Information regarding the medical consequences of a fall can inform both its severity and the clinical follow-up needed.
4. Injury after a fall is considered distinct from seeking and receiving medical attention.
5. Fall location can inform prevention efforts, including home safety; "some falls inside the home can be avoided through modifications such as removing a rug, better lighting etc."
6. Other details of the fall location can also inform home safety and prevention (eg, curb, stairs, and poor lighting).
7. Specifying whether falls occur because of factors related to multiple sclerosis or other factors (obstacle, etc) is important owing to the heterogeneity of fall triggers and of clinical responses.
8. The question "If you have a mobility aid, were you using it when you fell?" can remind patients to use the assistive device and can cue clinicians of the need to modify or change the current assistive device.
9. A falls log is provided to patients and shows the reported falls over time.

Closing the loop: real-time in-basket messaging

1. Enabling the reporting (patient) and ascertainment (clinician) of falls at regular intervals optimizes timeliness (vs periodic visits) while maintaining low burden (vs daily or "at time of fall"). If a severe fall is reported on the biweekly survey, an in-basket message to the electronic medical record alerts the care team in a manner integrated into the clinical workflow (Figure 6, #15).

XSL•FO
RenderX

Textbox 2. Multiple Sclerosis Falls InsightTrack clinical management dashboard integrated into the Epic electronic health record: key features. The numbers correspond to the red numbers in Figure 6, which denote key features mapping to the capability, opportunity, motivation, and behavior (COM-B) framework.

Longitudinal multiple sclerosis trajectory widget (visualizes patients' disease and medication trajectory over time with integrated normative ranges)

1. Ability to toggle through disability measures (eg, Expanded Disability Status Scale [EDSS] and Timed 25-Foot Walk)
2. Succinct overview of patient's longitudinal MS trajectory, including relapses, disability, medications, and normative data

Longitudinal falls widget (visualizes falls reported every 2 wk by the patient using their patient-facing app [Figure 5] data regarding date, time, and severity of each fall on 1 display)

1. Fall severity visualized by color shade (grading falls by severity considered important to trigger an alert to the care team and to inform type of clinical response)
2. Ability to include a way to visualize the falls log with falls over time
3. Estimated time of day of the fall can inform further interventions needed, including vision check, home safety evaluation, and medication review (especially for Parkinson disease)
4. Time of day visualized with colors for daytime (lighter: yellow) and nighttime (darker: blue) preferred by all stakeholders

Community resources widget (map automatically displays the patient's home community and allows for web-based identification of MS health care professionals in their community)

1. Automated display of MS professionals (physical therapist, occupational therapist, and talk therapist) in the patient's community, which reduces barriers for patient to identify local recourses once physical therapy or other referrals have been placed
2. Contact information and driving navigations between the patient's home and the resources automated, can be pasted into the patient's after-visit summary or the clinician's note

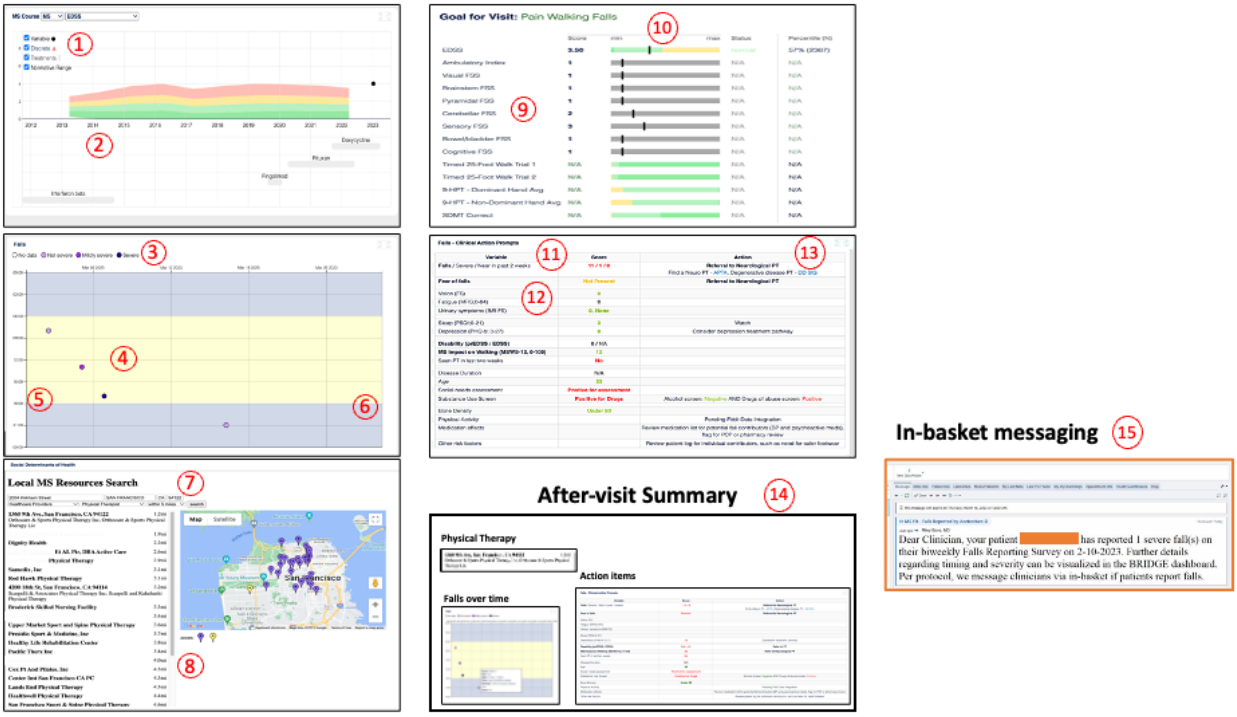
Cross-sectional widget (summary display with 2 tabs displaying clinical disability outcomes and patient-reported outcomes [PROs]; clinician can toggle between time points)

1. Clinic-based performance measures (walking speed, hand function, and cognition) and disability outcomes (EDSS with separate functional system scores) as well as PROs can inform a more global assessment of the patient at given time point
2. Color-coded normative ranges can provide rapid assessment of whether patient's given function is within "normal" range

Falls treatment– and action-prompt widget (tabulates core data needed for a comprehensive assessment of falls risk and prevention; for each category, patient's score is colored according to severity, and possible action prompts are displayed)

XSL•FO
RenderX

Figure 6. Multiple Sclerosis Falls InsightTrack (MS-FIT) clinical management dashboard, integrated into the Epic electronic health record (Epic Systems Corporation). This is the "MS-FIT BRIDGE version 3.0," which is viewable from Epic in the electronic health record at the time of the clinic visit. Key features mapping to user-generated perspectives and feedback and to the capability, opportunity, motivation, and behavior (COM-B) framework are denoted by a red number and described in Textbox 2.



Phase 4: Deliver

Altogether, 15 patients (10 with MS and 5 with PD) with an age range of 34 to 79 years and 6 MS clinicians with a clinical experience range of 2 to 22 years (Multimedia Appendix 1) launched the tool components *live* and provided feedback.

Patient-Facing UCSF MS-FIT

People With MS

Of the 15 patients, 10 (67%) had been diagnosed with MS; they had a mean age of 48.8 (SD 8.8; range 34-60) years, with disability level (EDSS score) ranging from 1.5 to 6.0 and a median disease duration of 14.5 (IQR 6.3-24; range 2-27) years. The feedback from people with MS was overwhelmingly positive (Table 1). Likability scores were all NPS≥100 (all promoters). The survey was found to be brief and clear. Patients appreciated the benefit of the closed-loop system and the overall impact on clinical encounters.

Patients With PD

Of the 15 patients, 5 (33%) had been diagnosed with PD; they had a mean age of 60.6 (SD 13.2; range 46-79) years, with a median disability level (Unified Parkinson's Disease Rating Scale score) of 31 (IQR 30.3-8.5; range 17-42) and a median disease duration of 4 (IQR 1.5-8.5, range 1-10) years. Overall, the NPS was found to be 0 (20%-20%, with 1/5, 20% detractor, 1/5, 20% promoter, and 3/5, 60% passive scores that trended toward promoters), indicating that patients with PD could be easily swayed to use MS-FIT. The mean scores on the Health-ITUES questions were all >4 (ie, agree or strongly agree), and only 1 score was <3 out of 5 (Table 2).

Qualitative insights from the interviews revealed that *falling*, *fear of falling*, and *thinking about falling* were "not at the top of their list," in contrast to people with MS. Nevertheless, patients with PD found the tracker "easy to fill out," and they "liked the idea of reporting falls and reporting if [they] experienced fear of falling." Patients with PD felt that it was important to have the ability to increase the font size (incorporated into MS-FIT patient survey v 2.0; Figure 5).

For future use in PD, patients reported that it would be important for ease of use and usability to have the ability to report motor vehicle accidents and specific PD symptoms as they relate to falls risk. Patients with PD also reported greater issues with using an iPad (motor or tremor issues).

XSL•FO
RenderX

Table 2. Scoring of the final University of California San Francisco Multiple Sclerosis Falls Tracker (REDCap [Research Electronic Data Capture]) by patients with Parkinson disease: usability, ease of use, and likability (n=5).

Health-ITUES[a]–based questions for usability, ease of use, and likability	Score, mean (SD)	Score <4 out of 5, n (%)
"It is useful to report if I've had any falls or near falls every 2 weeks"	4.40 (0.89)	1 (20)
"It is useful to have my survey answers sent to my care team"	4.20 (1.22)	1 (20)
"The survey asks about important topics"	4.00 (1.00)	2 (40)
"I am comfortable with my ability to complete the survey"	4.60 (0.89)	1 (20)
"I find the survey easy to use"	4.60 (0.89)	1 (20)
"I can easily remember how to access the survey through my email"	4.20 (1.10)	2 (40)
"I like the survey"	4.20 (1.10)	2 (40)

[a]Health-ITUES: Health IT Usability Evaluation Scale (scores range from 1=strongly disagree to 5=strongly agree).

MS-FIT Clinical Management Dashboard

Overall, the MS clinicians (n=6) rated the dashboard highly (NPS=16.67; Table 3):

> *I like that [the app] summarizes important clinical information in an easily digestible format, and the new widget that includes an MS [multiple sclerosis]-specific review of systems and actionable items seems like it will help ensure well-rounded MS care!* [Clinician 1]

With regard to reporting falls and near falls, the MS clinicians noted multiple benefits to aiding with patient care:

> *You can infer a lot from [fall data] in terms of disease activity, disease course, changes in a patient's life, their living setting, their support. If you see a jump in falls or the onset of falls in a patient who wasn't falling—it is worthy of clinical attention and needs to be addressed. It would give us an objective way to know if interventions are helping to reduce falls.* [Clinician 2]

> *Near falls are particularly underscreened, so any granularity on near falls would be helpful.* [Clinician 3]

> *For some patients, near falls may not be worth reporting—may just be part of life. But other patients it could make sense for. Any change from baseline has potential to be significant. Near falls can be [a] canary in a coal mine.* [Clinician 5]

Table 3. Scoring of the final University of California San Francisco Multiple Sclerosis Falls BRIDGE dashboard by multiple sclerosis clinicians: usability, ease of use, and likability.

Health-ITUES[a]–based questions for usability, ease of use, and likability	Score, mean (SD)	Score <4 out of 5, n (%)
"The information that appears in BRIDGE is useful to me."	4.80 (0.41)	0 (0)
"It is useful to be updated on my patient's significant fall activity between appointments."	4.50 (0.84)	1 (17)
"I find BRIDGE easy to use."	4.20 (0.75)	1 (20)
"I can always remember how to access BRIDGE."	4.00 (1.10)	1 (20)
"I like BRIDGE."	4.50 (0.55)	0 (0)

[a]Health-ITUES: Health IT Usability Evaluation Scale (scores range from 1=strongly disagree to 5=strongly agree).

Discussion

Principal Findings

To our knowledge, this is the first tool designed using the HCD framework, anchored in the COM-B approach to behavior change, and capable of delivering relevant information at the point of care in line with the *5 rights* with the aim of preventing falls in people with MS. Other apps have been developed, although the focus has mainly been on 1 component of falls (eg, evaluating falls risk [51]) at a time. In addition, many large-scale clinical research projects, such as those conducted at the Stanford Center for Digital Health and the Remote Assessment of Disease and Relapse–Central Nervous System program, are exploring applications of wearable data. However, most of the collected wearable data remain inaccessible for visualization or integration within a clinic's EHR. This limitation can impede the effective use of PGD by clinicians and compromise patient-physician collaboration related to PGD [52]. MS-FIT fills a critical gap in multimodal closed-loop self-management apps for falls monitoring and prevention.

Through extensive stakeholder engagement, MS-FIT offers novel aspects of customization, generalizability, and scalability, integrating multiple data streams relevant to reducing falls. It provides rapid personalized in-basket notifications, limited to severe falls, and digitally displays PGD through the EHR, increasing the likelihood of adoption by patients and clinicians.

Designed in collaboration with patients and clinicians, MS-FIT has emerged as a well-received closed-loop tool for tracking falls and reducing falls risk in individuals with MS. Patients liked its brevity, simplicity, and overall utility, recognizing its

XSL•FO
RenderX

potential to enhance clinical discussions. The utility of between-visit reporting and contextualized information for identifying modifiable falls risks was acknowledged by both patients and clinicians. The trial phase aims to validate its low-burden design in practice. Clinicians welcomed the closed-loop system, foreseeing proactive interventions and streamlined implementation. Longitudinal falls visualization, incorporating time and severity, along with clinician prompts targeting MS symptoms and medication effects, was favored for its ability to capture often overlooked components during regular visits.

Another noteworthy finding was the minimal number of interviews required to attain thematic saturation in our initial discover phase, indicating that some clear guidance for potentially high-value initial design features was achieved with a minimal sample size. This could be due to the fact that MS-FIT was based on an initial prototype developed during a prediscover phase using patient and clinician feedback. It could also be attributed to homogeneous samples of study participants consulted throughout the discover and design phases. Overall design efficiency was likely aided by the experience and regular input of interprofessional teams.

The ongoing process involves testing MS-FIT in a prospective longitudinal study in a cohort of 100 adults with MS over 12 months. The primary objectives of this larger study include assessing the adoption rate of the tool, evaluating the level of sustained use of the tool, monitoring adherence to falls reporting, and assessing study retention over the 12-month period. Secondary and exploratory analyses will center around the prediction of adoption, sustained use, adherence to action prompts, and study retention. To determine effectiveness, the study will compare in-study falls with a prior falls data set (Fitbit remote monitoring in MS) [41], and patient satisfaction will be assessed during an exit interview.

Scalability

Our approach, characterized by the selection of key technological and clinical features, allows for the scalability and generalizability of the tool's modular infrastructure to various symptoms, conditions, and clinical settings for other high fall-risk diseases as well as other symptoms within MS (eg, bladder dysfunction). Technological factors for scalability include (1) high-quality, widely shareable static visualizations; and (2) optimized industry standards for code sharing with clinicians in other health care settings, such as other MS centers using Epic EHR. However, successful integration into other health systems depends on the internal governance and motivation within each system.

Limitations

All interviews were conducted remotely, using the UCSF Zoom video platform, which may have biased the patient stakeholders to people who are technologically literate and have access to the internet. However, 92% of people in the United States have access to the internet [53], and given that MS-FIT is an app, users (patients or caregivers) are expected to possess a certain level of technical proficiency. Only clinicians at UCSF and patients seen by this (broad) group of clinicians were interviewed; therefore, we may have missed important feedback from a wider cohort of users. Although HCD is favored for user-driven eHealth innovations, certain limitations exist [32], including a narrow focus; thus, exploring alternatives such as value-sensitive design, citizen science, and more-than-human design could enhance inclusivity and impact within eHealth innovation [54]. Finally, having the interviewer serve as the primary coder could have introduced bias into the qualitative analysis process. Stakeholder advisory group engagement in the coding process was an effort to reduce any potential bias.

Conclusions

MS-FIT delivers relevant data to clinicians through an embedded window within the EHR, following the *5 rights* approach. By using MS-FIT for data processing and algorithms, we reduce clinician burden while enhancing care. Our innovation extends to enabling and integrating real-world PGD as well as clinical and community-level factors, providing actionable information to empower self-care and addressing the impact of falls in people with MS. Our preliminary data indicate that this tool and design extend beyond MS and can be applied to other conditions associated with falls as well as the fear of falls and their associated consequences. To test the feasibility and effectiveness of the app, a clinical trial is ongoing (University of California San Francisco Clinical Trials identifier: NCT05837949).

Acknowledgments

This study was funded by the National Library of Medicine award (R01LM013396; principal investigator: RB). VJB is supported by the National Multiple Sclerosis Society Career Transition Award. RB is funded by the National Multiple Sclerosis Society Harry Weaver Scholar Award.

Data Availability

The data sets generated and analyzed during this study are available from the corresponding author on reasonable request.

Conflicts of Interest

VJB is funded by the National Multiple Sclerosis Society Career Transition Award. CYG provides medical consulting for EMD Serono, Genentech, and Horizon Therapeutics. JMG receives research support to University of California San Francisco from Genentech, Hoffmann-La Roche, Vigil Neuroscience, and consulting for Arialys. EB reports research funding from the Michael J Fox Foundation, the Gateway Foundation for Brain Research, the National Institutes of Health, and Biogen Inc. EB has also

XSL•FO
RenderX

received honoraria for his work as neurology section editor of NEJM Knowledge+ and consulting fees from Rune Labs, Inc. IS is the cofounder of Open mHealth; general assembly member for The Commons Project; serves on the medical advisory board for 98point6 Technologies; is cofounder, board director, and consultant for Vivli; is on the scientific advisory board for Myovant; and holds stocks in Myia. RB is a recipient of funding from the National Multiple Sclerosis Society Harry Weaver Award, the National Institutes of Health, DOD, and NSF, as well as Biogen, Novartis, and Roche Genentech. She has received personal fees for consulting from Alexion, EMD Serono, Horizon, Jansen, Genzyme Sanofi, and TG Therapeutics. All other authors declare no other conflicts of interest.

Multimedia Appendix 1
Demographic information for each round of interviews that led to the development of Multiple Sclerosis Falls InsightTrack.
[DOCX File , 16 KB - humanfactors_v11i1e49331_app1.docx]

Multimedia Appendix 2
Overview of findings from interviews with clinicians and participants with multiple sclerosis (MS), highlighting areas that block or boost patient and clinician behavior change with regard to falls and falls prevention. The table indicates whether intervention solution features were incorporated into the Multiple Sclerosis Falls InsightTrack (MS-FIT) patient survey, the clinician dashboard, or both.
[DOCX File , 33 KB - humanfactors_v11i1e49331_app2.docx]

References

1. Coote S, Comber L, Quinn G, Santoyo-Medina C, Kalron A, Gunn H. Falls in people with multiple sclerosis: risk identification, intervention, and future directions. Int J MS Care 2020;22(6):247-255 [FREE Full text] [doi: 10.7224/1537-2073.2020-014] [Medline: 33424479]
2. Cattaneo D, De Nuzzo C, Fascia T, Macalli M, Pisoni I, Cardini R. Risks of falls in subjects with multiple sclerosis. Arch Phys Med Rehabil 2002 Jun;83(6):864-867. [doi: 10.1053/apmr.2002.32825] [Medline: 12048669]
3. Tajali S, Shaterzadeh-Yazdi MJ, Negahban H, van Dieën JH, Mehravar M, Majdinasab N, et al. Predicting falls among patients with multiple sclerosis: comparison of patient-reported outcomes and performance-based measures of lower extremity functions. Mult Scler Relat Disord 2017 Oct;17:69-74. [doi: 10.1016/j.msard.2017.06.014] [Medline: 29055478]
4. Cameron MH, Nilsagard Y. Balance, gait, and falls in multiple sclerosis. Handb Clin Neurol 2018;159:237-250. [doi: 10.1016/B978-0-444-63916-5.00015-X] [Medline: 30482317]
5. Haddad YK, Bergen G, Florence CS. Estimating the economic burden related to older adult falls by state. J Public Health Manag Pract 2019;25(2):E17-E24 [FREE Full text] [doi: 10.1097/PHH.0000000000000816] [Medline: 29757813]
6. Peterson EW, Cho CC, von Koch L, Finlayson ML. Injurious falls among middle aged and older adults with multiple sclerosis. Arch Phys Med Rehabil 2008 Jun;89(6):1031-1037. [doi: 10.1016/j.apmr.2007.10.043] [Medline: 18503796]
7. Kalron A, Aloni R, Givon U, Menascu S. Fear of falling, not falls, impacts leisure-time physical activity in people with multiple sclerosis. Gait Posture 2018 Sep;65:33-38. [doi: 10.1016/j.gaitpost.2018.06.174] [Medline: 30558943]
8. Choi K, Jeon GS, Cho SI. Prospective study on the impact of fear of falling on functional decline among community dwelling elderly women. Int J Environ Res Public Health 2017 Apr 27;14(5):469 [FREE Full text] [doi: 10.3390/ijerph14050469] [Medline: 28448461]
9. Schoene D, Heller C, Aung YN, Sieber CC, Kemmler W, Freiberger E. A systematic review on the influence of fear of falling on quality of life in older people: is there a role for falls? Clin Interv Aging 2019 Apr 24;14:701-719 [FREE Full text] [doi: 10.2147/CIA.S197857] [Medline: 31190764]
10. Coote S, Sosnoff JJ, Gunn H. Fall incidence as the primary outcome in multiple sclerosis falls-prevention trials: recommendation from the international MS falls prevention research network. Int J MS Care 2014;16(4):178-184 [FREE Full text] [doi: 10.7224/1537-2073.2014-059] [Medline: 25694776]
11. Gunn H, Markevics S, Haas B, Marsden J, Freeman J. Systematic review: the effectiveness of interventions to reduce falls and improve balance in adults with multiple sclerosis. Arch Phys Med Rehabil 2015 Oct;96(10):1898-1912. [doi: 10.1016/j.apmr.2015.05.018] [Medline: 26070975]
12. Lee HS, Park YJ, Park SW. The effects of virtual reality training on function in chronic stroke patients: a systematic review and meta-analysis. Biomed Res Int 2019 Jun 18;2019:7595639 [FREE Full text] [doi: 10.1155/2019/7595639] [Medline: 31317037]
13. Matsuda PN, Shumway-Cook A, Bamer AM, Johnson SL, Amtmann D, Kraft GH. Falls in multiple sclerosis. PM R 2011 Jul 21;3(7):624-32; quiz 632. [doi: 10.1016/j.pmrj.2011.04.015] [Medline: 21777861]
14. Mazuz K, Biswas S, Lindner U. Developing self-management application of fall prevention among older adults: a content and usability evaluation. Front Digit Health 2020 Sep 2;2:11 [FREE Full text] [doi: 10.3389/fdgth.2020.00011] [Medline: 34713024]
15. Rajagopalan R, Litvan I, Jung TP. Fall prediction and prevention systems: recent trends, challenges, and future research directions. Sensors (Basel) 2017 Nov 01;17(11):2509 [FREE Full text] [doi: 10.3390/s17112509] [Medline: 29104256]

XSL•FO
RenderX

16. Bet P, Castro PC, Ponti MA. Fall detection and fall risk assessment in older person using wearable sensors: a systematic review. Int J Med Inform 2019 Oct;130:103946. [doi: 10.1016/j.ijmedinf.2019.08.006] [Medline: 31450081]
17. Klaren RE, Motl RW, Dlugonski D, Sandroff BM, Pilutti LA. Objectively quantified physical activity in persons with multiple sclerosis. Arch Phys Med Rehabil 2013 Dec;94(12):2342-2348. [doi: 10.1016/j.apmr.2013.07.011] [Medline: 23906692]
18. Gelfand JM. Multiple sclerosis: diagnosis, differential diagnosis, and clinical presentation. Handb Clin Neurol 2014;122:269-290. [doi: 10.1016/B978-0-444-52001-2.00011-X] [Medline: 24507522]
19. Hennessey A, Robertson NP, Swingler R, Compston DA. Urinary, faecal and sexual dysfunction in patients with multiple sclerosis. J Neurol 1999 Nov;246(11):1027-1032. [doi: 10.1007/s004150050508] [Medline: 10631634]
20. Motl RW, McAuley E, Dlugonski D. Reactivity in baseline accelerometer data from a physical activity behavioral intervention. Health Psychol 2012 Mar;31(2):172-175. [doi: 10.1037/a0025965] [Medline: 22023436]
21. Motl RW, Pilutti LA, Learmonth YC, Goldman MD, Brown T. Clinical importance of steps taken per day among persons with multiple sclerosis. PLoS One 2013 Sep 4;8(9):e73247 [FREE Full text] [doi: 10.1371/journal.pone.0073247] [Medline: 24023843]
22. Sandroff BM, Dlugonski D, Weikert M, Suh Y, Balantrapu S, Motl RW. Physical activity and multiple sclerosis: new insights regarding inactivity. Acta Neurol Scand 2012 Oct 03;126(4):256-262. [doi: 10.1111/j.1600-0404.2011.01634.x] [Medline: 22211941]
23. Adams SD, Kouzani AZ, Tye SJ, Bennet KE, Berk M. An investigation into closed-loop treatment of neurological disorders based on sensing mitochondrial dysfunction. J Neuroeng Rehabil 2018 Feb 13;15(1):8 [FREE Full text] [doi: 10.1186/s12984-018-0349-z] [Medline: 29439744]
24. Wadwa RP, Reed ZW, Buckingham BA, DeBoer MD, Ekhlaspour L, Forlenza GP, et al. Trial of hybrid closed-loop control in young children with type 1 diabetes. N Engl J Med 2023 Mar 16;388(11):991-1001 [FREE Full text] [doi: 10.1056/NEJMoa2210834] [Medline: 36920756]
25. Lo MC, Widge AS. Closed-loop neuromodulation systems: next-generation treatments for psychiatric illness. Int Rev Psychiatry 2017 Apr;29(2):191-204 [FREE Full text] [doi: 10.1080/09540261.2017.1282438] [Medline: 28523978]
26. Quinn G, Comber L, Galvin R, Coote S. The ability of clinical balance measures to identify falls risk in multiple sclerosis: a systematic review and meta-analysis. Clin Rehabil 2018 May 20;32(5):571-582. [doi: 10.1177/0269215517748714] [Medline: 29260583]
27. Quinn G, Comber L, McGuigan C, Galvin R, Coote S. Discriminative ability and clinical utility of the Timed Up and Go (TUG) in identifying falls risk in people with multiple sclerosis: a prospective cohort study. Clin Rehabil 2019 Feb 14;33(2):317-326. [doi: 10.1177/0269215518793481] [Medline: 30103642]
28. Block V, Pitsch E, Gopal A, Zhao C, Pletcher MJ, Marcus GM, et al. Identifying falling risk factors remotely in multiple sclerosis (2597). Neurology 2021;96(15 Supplement):2597.
29. Sebastião E, Learmonth YC, Motl RW. Lower physical activity in persons with multiple sclerosis at increased fall risk: a cross-sectional study. Am J Phys Med Rehabil 2017 May;96(5):357-361. [doi: 10.1097/PHM.0000000000000581] [Medline: 28415071]
30. Campbell JD. The five "rights" of clinical decision support. J AHIMA 2013 Oct;84(10):42-7; quiz 48. [Medline: 24245088]
31. Brown T. Design thinking. Harv Bus Rev 2008 Jun;86(6):84-92, 141. [Medline: 18605031]
32. Göttgens I, Oertelt-Prigione S. The application of human-centered design approaches in health research and innovation: a narrative review of current practices. JMIR Mhealth Uhealth 2021 Dec 06;9(12):e28102 [FREE Full text] [doi: 10.2196/28102] [Medline: 34874893]
33. Brown W3, Yen PY, Rojas M, Schnall R. Assessment of the health IT usability evaluation model (Health-ITUEM) for evaluating mobile health (mHealth) technology. J Biomed Inform 2013 Dec;46(6):1080-1087 [FREE Full text] [doi: 10.1016/j.jbi.2013.08.001] [Medline: 23973872]
34. Design kit: the human-centered design toolkit. IDEO. 2015 Jun 11. URL: https://www.ideo.com/post/design-kit [accessed 2023-07-07]
35. Matheson GO, Pacione C, Shultz RK, Klügl M. Leveraging human-centered design in chronic disease prevention. Am J Prev Med 2015 Apr;48(4):472-479. [doi: 10.1016/j.amepre.2014.10.014] [Medline: 25700655]
36. The design process: what is the double diamond? Design Council. URL: https://www.designcouncil.org.uk/news-opinion/design-process-what-double-diamond [accessed 2023-07-24]
37. Brown EG, Schleimer E, Bledsoe IO, Rowles W, Miller NA, Sanders SJ, et al. Enhancing clinical information display to improve patient encounters: human-centered design and evaluation of the Parkinson disease-BRIDGE platform. JMIR Hum Factors 2022 May 06;9(2):e33967 [FREE Full text] [doi: 10.2196/33967] [Medline: 35522472]
38. Bove R, Schleimer E, Sukhanov P, Gilson M, Law SM, Barnecut A, et al. Building a precision medicine delivery platform for clinics: the University of California, San Francisco, BRIDGE experience. J Med Internet Res 2022 Feb 15;24(2):e34560 [FREE Full text] [doi: 10.2196/34560] [Medline: 35166689]
39. Block VJ, Gopal A, Rowles W, Yueh C, Gelfand JM, Bove R. CoachMS, an innovative closed-loop, interdisciplinary platform to monitor and proactively treat MS symptoms: a pilot study. Mult Scler J Exp Transl Clin 2021 Feb 02;7(1):2055217321988937 [FREE Full text] [doi: 10.1177/2055217321988937] [Medline: 33796329]

XSL•FO
RenderX

40. Gopal A, Gelfand JM, Bove R, Block VJ. Fall assessment and monitoring in people with multiple sclerosis: a practical evidence-based review for clinicians. Neurol Clin Pract 2023 Oct;13(5):e200184 [FREE Full text] [doi: 10.1212/CPJ.0000000000200184] [Medline: 37720138]
41. Block VJ, Pitsch EA, Gopal A, Zhao C, Pletcher MJ, Marcus GM, et al. Identifying falls remotely in people with multiple sclerosis. J Neurol 2022 Apr;269(4):1889-1898 [FREE Full text] [doi: 10.1007/s00415-021-10743-y] [Medline: 34405267]
42. Bove R, Bruce CA, Lunders CK, Pearce JR, Liu J, Schleimer E, et al. Electronic health record technology designed for the clinical encounter: MS NeuroShare. Neurol Clin Pract 2021 Aug;11(4):318-326 [FREE Full text] [doi: 10.1212/CPJ.0000000000000986] [Medline: 34484932]
43. Schleimer E, Pearce J, Barnecut A, Rowles W, Lizee A, Klein A, et al. A precision medicine tool for patients with multiple sclerosis (the open MS BioScreen): human-centered design and development. J Med Internet Res 2020 Jul 06;22(7):e15605 [FREE Full text] [doi: 10.2196/15605] [Medline: 32628124]
44. Schnall R, Cho H, Liu J. Health information technology usability evaluation scale (health-ITUES) for usability assessment of mobile health technology: validation study. JMIR Mhealth Uhealth 2018 Jan 05;6(1):e4 [FREE Full text] [doi: 10.2196/mhealth.8851] [Medline: 29305343]
45. Mathews SC, McShea MJ, Hanley CL, Ravitz A, Labrique AB, Cohen AB. Digital health: a path to validation. NPJ Digit Med 2019 May 13;2:38 [FREE Full text] [doi: 10.1038/s41746-019-0111-3] [Medline: 31304384]
46. Frankfort-Nachmias C, Nachmias D. Research Methods in the Social Sciences 7th edition. New York, NY: Worth Publishers; Mar 6, 2007.
47. Designing consumer health IT: a guide for developers and systems designers. Agency for Healthcare Research and Quality. 2012 Sep. URL: https://digital.ahrq.gov/sites/default/files/docs/page/designing-consumer-health-it-a-guide-for-developers-and-systems-designers.pdf [accessed 2023-07-01]
48. Michie S, van Stralen MM, West R. The behaviour change wheel: a new method for characterising and designing behaviour change interventions. Implement Sci 2011 Apr 23;6:42 [FREE Full text] [doi: 10.1186/1748-5908-6-42] [Medline: 21513547]
49. Bucher A. Engaged: Designing for Behavior Change. Brooklyn, NY: Rosenfeld Media; Mar 03, 2020.
50. The patient education materials assessment tool (PEMAT) and user's guide. Agency for Healthcare Research and Quality. URL: https://www.ahrq.gov/health-literacy/patient-education/pemat.html [accessed 2023-04-09]
51. Hsieh K, Fanning J, Frechette M, Sosnoff J. Usability of a fall risk mHealth app for people with multiple sclerosis: mixed methods study. JMIR Hum Factors 2021 Mar 22;8(1):e25604 [FREE Full text] [doi: 10.2196/25604] [Medline: 33749609]
52. Lordon RJ, Mikles SP, Kneale L, Evans HL, Munson SA, Backonja U, et al. How patient-generated health data and patient-reported outcomes affect patient-clinician relationships: a systematic review. Health Informatics J 2020 Dec;26(4):2689-2706 [FREE Full text] [doi: 10.1177/1460458220928184] [Medline: 32567460]
53. Percentage of population using the internet in the United States from 2000 to 2023. Statista. URL: https://www.statista.com/statistics/209117/us-internet-penetration/#:~:text=As%20of%202023%2C%20approximately%2092,internet%20users%20in%20the%20country [accessed 2023-07-01]
54. van Velsen L, Ludden G, Grünloh C. The limitations of user-and human-centered design in an eHealth context and how to move beyond them. J Med Internet Res 2022 Oct 05;24(10):e37341 [FREE Full text] [doi: 10.2196/37341] [Medline: 36197718]

Abbreviations

API: application programming interface
BCW: Behavioral Change Wheel
COM-B: capability, opportunity, motivation, and behavior
CSS: cascading style sheets
EHR: electronic health record
FHIR: Fast Healthcare Interoperability Resources
HCD: human-centered design
Health-ITUEM: Health IT Usability Evaluation Model
Health-ITUES: Health IT Usability Evaluation Scale
MS: multiple sclerosis
MS-FIT: Multiple Sclerosis Falls InsightTrack
PD: Parkinson disease
PGD: patient-generated data
PT: physical therapy
REDCap: Research Electronic Data Capture
SMART: Substitutable Medical Apps and Reusable Technologies
UCSF: University of California San Francisco

XSL•FO
RenderX

Edited by A Choudhury; submitted 26.05.23; peer-reviewed by H Gooch, E Prochaska; comments to author 27.07.23; revised version received 14.08.23; accepted 19.10.23; published 11.01.24.

Please cite as:
Block VJ, Koshal K, Wijangco J, Miller N, Sara N, Henderson K, Reihm J, Gopal A, Mohan SD, Gelfand JM, Guo CY, Oommen L, Nylander A, Rowson JA, Brown E, Sanders S, Rankin K, Lyles CR, Sim I, Bove R
A Closed-Loop Falls Monitoring and Prevention App for Multiple Sclerosis Clinical Practice: Human-Centered Design of the Multiple Sclerosis Falls InsightTrack
JMIR Hum Factors 2024;11:e49331
URL: https://humanfactors.jmir.org/2024/1/e49331
doi:10.2196/49331
PMID:38206662

©Valerie J Block, Kanishka Koshal, Jaeleene Wijangco, Nicolette Miller, Narender Sara, Kyra Henderson, Jennifer Reihm, Arpita Gopal, Sonam D Mohan, Jeffrey M Gelfand, Chu-Yueh Guo, Lauren Oommen, Alyssa Nylander, James A Rowson, Ethan Brown, Stephen Sanders, Katherine Rankin, Courtney R Lyles, Ida Sim, Riley Bove. Originally published in JMIR Human Factors (https://humanfactors.jmir.org), 11.01.2024. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on https://humanfactors.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Digital Triage Tools for Sexually Transmitted Infection Testing Compared With General Practitioners' Advice: Vignette-Based Qualitative Study With Interviews Among General Practitioners

Kyma Schnoor[1,2], MSc; Anke Versluis[1,2], PhD; Niels H Chavannes[1,2], MD, PhD; Esther P W A Talboom-Kamp[1,2,3], MBA, MD, PhD

[1]Public Health and Primary Care, Leiden University Medical Center, Leiden, Netherlands
[2]National eHealth Living Lab, Leiden University Medical Center, Leiden, Netherlands
[3]Zuyderland, Sittard-Geleen, Netherlands

Corresponding Author:
Kyma Schnoor, MSc
Public Health and Primary Care
Leiden University Medical Center
Hippocratespad 21
Leiden, 2333 ZA
Netherlands
Phone: 31 71526 8433
Email: k.schnoor@lumc.nl

Abstract

Background: Digital triage tools for sexually transmitted infection (STI) testing can potentially be used as a substitute for the triage that general practitioners (GPs) perform to lower their work pressure. The studied tool is based on medical guidelines. The same guidelines support GPs' decision-making process. However, research has shown that GPs make decisions from a holistic perspective and, therefore, do not always adhere to those guidelines. To have a high-quality digital triage tool that results in an efficient care process, it is important to learn more about GPs' decision-making process.

Objective: The first objective was to identify whether the advice of the studied digital triage tool aligned with GPs' daily medical practice. The second objective was to learn which factors influence GPs' decisions regarding referral for diagnostic testing. In addition, this study provides insights into GPs' decision-making process.

Methods: A qualitative vignette-based study using semistructured interviews was conducted. In total, 6 vignettes representing patient cases were discussed with the participants (GPs). The participants needed to think aloud whether they would advise an STI test for the patient and why. A thematic analysis was conducted on the transcripts of the interviews. The vignette patient cases were also passed through the digital triage tool, resulting in advice to test or not for an STI. A comparison was made between the advice of the tool and that of the participants.

Results: In total, 10 interviews were conducted. Participants (GPs) had a mean age of 48.30 (SD 11.88) years. For 3 vignettes, the advice of the digital triage tool and of all participants was the same. In those vignettes, the patients' risk factors were sufficiently clear for the participants to advise the same as the digital tool. For 3 vignettes, the advice of the digital tool differed from that of the participants. Patient-related factors that influenced the participants' decision-making process were the patient's anxiety, young age, and willingness to be tested. Participants would test at a lower threshold than the triage tool because of those factors. Sometimes, participants wanted more information than was provided in the vignette or would like to conduct a physical examination. These elements were not part of the digital triage tool.

Conclusions: The advice to conduct a diagnostic STI test differed between a digital triage tool and GPs. The digital triage tool considered only medical guidelines, whereas GPs were open to discussion reasoning from a holistic perspective. The GPs' decision-making process was influenced by patients' anxiety, willingness to be tested, and age. On the basis of these results, we believe that the digital triage tool for STI testing could support GPs and even replace consultations in the future. Further research must substantiate how this can be done safely.

(JMIR Hum Factors 2024;11:e49221) doi:10.2196/49221

XSL•FO
RenderX

KEYWORDS

eHealth; digital triage tool; sexually transmitted infection; STI; human immunodeficiency virus; general practitioners; GPs decision-making; digital health; diagnostic; sexually transmitted disease; STD; sexually transmitted; sexual transmission; triage; artificial intelligence; HIV; diagnostics; diagnosis; vignette; vignettes; interview; interviews; best practice; best practices; thematic analysis; referral; medical advice

Introduction

Background

The use of eHealth, health services delivered through the internet or related technologies, is increasing, especially since the COVID-19 pandemic [1,2]. The COVID-19 pandemic has shed light on the crucial role of digitization in health care [2]. An important and promising element of digitization in health care are digital triage tools consisting of a questionnaire for patients to identify the risk of a medical problem. These tools use a digital questionnaire typically administered by a health care professional, and an algorithm based on a medical decision tree generates automatic advice for follow-up, for example, a web-based symptom checker. In this paper, we discuss a digital triage tool that advises whether a specific diagnostic test for a specific combination of symptoms is necessary. This specific digital triage tool is based on Dutch medical guidelines.

Such a digital triage tool for different problems and symptoms could be an efficient and accessible method for citizens with medical questions. In addition, this digital triage tool could possibly lower the workload of general practitioners (GPs) as it can replace the triage that health care professionals would do themselves [3]. However, it is important that triage leads to responsible and appropriate care given the situation. Digital triage tools should not result in “over-triage” or “under-triage” [4]. Overtriage is when a patient is advised to undergo a medical treatment or diagnostic test when they do not have an (urgent) medical problem [4]. Undertriage is when a patient is told that they do not have an (urgent) medical problem when they do, with the advice that a diagnostic test or medical treatment is not necessary [4]. It is important to know whether the digital triage tool for diagnostic tests is in line with daily medical practice to maximize its validity.

In daily practice at GPs’ offices, medical guidelines are used to support their decision-making. GPs following guidelines has been an important research subject into the decision-making process of GPs in dermatology has shown that GPs do not always adhere to medical guidelines [5]. For example, concerns about the patient or the relationship between the GP and the patient were sometimes part of the decision-making process [5]. Furthermore, a meta-synthesis of qualitative studies identified GPs’ attitudes toward and experiences with clinical guidelines [6]. First, this study showed that GPs experience tension between their own experiences and the guidelines they must adhere to as guidelines do not consider personal circumstances. Second, GPs are afraid of missing a patient diagnosis. Third, GPs experience that the guidelines do not always fit with patients’ needs, and therefore, GPs act differently from what the guidelines instruct them to do. Earlier reviews have revealed other factors that play a role in the decision-making process of GPs in referrals for diagnostic tests [7-9]. These are, among others, demographic and nonclinical factors such as patient characteristics (eg, age, sex, and social class [8]). In addition, the patient’s quality of life and wishes are nonclinical factors that influence the decision-making process of the GP [7]. Not all those factors are included in medical guidelines and, consequently, in digital triage. All these factors clearly show that the GP makes decisions from a holistic perspective, which makes it even more interesting and important to critically consider decision-making using digital tools from the perspective of the GP. Regarding diagnostic testing, to our knowledge, our study is the first one that compares the advice of GPs with that of a web-based tool. At the same time, this study identifies what factors influence a GP’s decision-making process for a diagnostic test.

Objectives

If a digital triage tool is of high quality and the patient is adequately advised, a consultation with the GP could be avoided, resulting in an efficient care process for the patient. The GP can also be supported in the hectic daily workload as the patient uses the tool independently [9]. The first objective of this study was to identify whether the advice of the studied digital triage tool aligned with the daily medical practice of the GP. The second objective was to learn which factors influenced the GP’s decision regarding a referral for diagnostic testing. In addition, this research provides insights into the GP’s decision-making process and whether factors are possibly missing from a digital triage tool. As a starting point, we investigated these research questions for sexually transmitted infection (STI) triage as the medical guidelines are straightforward (eg, clear risk factors and answer categories). Much research has been conducted on digital applications for STI testing, such as websites in which tests can be ordered, with positive feedback from patients about their usability [10]. Moreover, research has shown that a digital triage tool can potentially lower the threshold for STI testing [10] as this problem can be associated with feelings of shame [11]. To answer the research questions, a vignette-based qualitative study was conducted based on different STI-related patient cases [12].

Methods

Study Design and Participants

A qualitative vignette study was conducted using semistructured interviews with GPs as participants. Data saturation was expected after 10 interviews [13]. There were no specific exclusion criteria. GPs in training, practicing, or retired (for ≤5 y) could participate. In the interviews, the participants were presented with different patient vignettes (see the *Materials* section for details). After each vignette, the participants were asked about their clinical decision regarding STI diagnostic testing and to describe their thinking and decision-making process. This approach is called the “Think Aloud” method,

XSL•FO
RenderX

which allows for a description of how information is structured during a problem-solving task [14]. In addition, it provides rich data for analysis [15].

Ethical Considerations

This study was declared not to fall within the scope of the Dutch Medical Research Involving Human Subjects Act by the departmental ethics committee of the Leiden University Medical Center (reference 22-3002).

Materials

A vignette is a short hypothetical description of a patient representing a standardized combination of specific characteristics [16]. Vignettes made it possible to present patients with the same characteristics to every participant (eg, complaints, relationship status, and age) and, in this way, minimize variations between patients, which is not possible in real life. In this study, the vignettes were based on different aspects of the Dutch medical guidelines for STI testing [17]. In the medical guidelines, different aspects are taken into account to calculate the risk of an STI, such as endemic areas, unsafe sex, and different complaints. The following factors were incorporated into the vignettes: age, gender, sexuality, relationship status, employment (eg, full-time job or student), history of unsafe sex and how long ago it took place, number of sexual partners, frequency of unsafe sex, frequent GP visits, symptoms, and ethnicity. Some of these factors are not in the guidelines but were included to research whether they influenced the decision-making process of the GP (eg, situation and if the GP was visited often by that patient). In addition, the vignettes were designed in such a way that they would lead to advice from participants to undergo a diagnostic test for STIs or not. In total, 6 different vignettes were created and used (Multimedia Appendix 1). In Textbox 1, a short description of the vignettes is provided. The Dutch vignettes were designed with a GP and checked by another GP. An example of a translated vignette can be found in Textbox 2.

Textbox 1. Short description of the vignettes.

Vignette 1

- Woman, aged 20 years, from Spain, student, had unsafe sex multiple times >3 weeks ago, itching of the vagina, does not visit her general practitioner (GP) often

Vignette 2

- Man, aged 26 years, plumber, steady relationship, has irritation at the urethra and sensitivity when urinating, visits GP often

Vignette 3

- Woman, aged 17 years, high school student, had unsafe sex <3 weeks ago with no complaints, the first time she comes to the practice

Vignette 4

- Man, aged 24 years, has a relationship with a man, his partner has sexual contact with other men, has difficulty urinating

Vignette 5

- Woman, aged 45 years, has a steady relationship but thinks her partner cheated 6 months ago, has contact bleeding, visits the GP often

Vignette 6

- Woman, aged 35 years, has a steady relationship, comes from Surinam, has a burning sensation when urinating, visits her GP often

Textbox 2. Vignette 1 translated from Dutch to English.

- Mrs A is aged 20 years and studies in the Netherlands but comes from Spain originally. She has not visited you at the practice often. She is not in a committed relationship and has had unprotected sex several times in the past 6 months for more than 3 weeks. She experiences vaginal discharge and itching and irritation in her vagina. She wonders whether she might have a sexually transmitted infection.

Procedure

Participants were recruited via a LinkedIn post that included the email address of the researcher. Interested participants were instructed to send an email if they wanted to take part. In addition, participants were emailed from the network of the researchers, and the GPs could reply to the email if they wanted to participate. Interested participants were sent information and the informed consent form. In addition, different data and time points were included in the interviews, which could be face-to-face or digital (based on the preference of the participant). Participants had the right to withdraw at any time.

An interview protocol guided the semistructured interviews (Multimedia Appendix 2). All interviews were audio recorded. Each interview started with a short explanation of the study. The first vignette was then read out loud to the participant. They were asked whether they would advise undergoing diagnostic tests for STIs. Next, they were asked to share their reasoning process. These 2 steps were repeated for each vignette (ie, 6 in total). The first interviews were conducted with both interviewers present (KS and Fleur Rekveld), and KS was the

XSL•FO
RenderX

lead. The other interviews were conducted by KS, Fleur Rekveld, or both.

Service: Digital Triage Tool

The digital triage tool was developed by a Dutch diagnostic center [18] based on a decision tree with Dutch medical guidelines [17]. The digital triage tool was developed in cocreation with GPs and clinical chemists. A Dutch academic knowledge center assessed the digital triage [19]. During triage, users first go through a series of questions. Their answers determine what question they have to answer next and, in the end, what advice is given. For example, the first question is "Did you have unsafe sex?" If the answer is "no," the advice is not to be tested. If the answer is "yes," a follow-up question appears: *what is your gender?* Gender is asked about as differences in gender result in different advice (eg, for women users who are advised to undergo a chlamydia test, it means that the service could advise doing a vaginal swab). Ultimately, the digital triage tool advises whether a diagnostic test for STIs is necessary and, if yes, which one (eg, chlamydia, gonorrhea, or HIV). The digital triage tool is now used in 2 digital services of the diagnostic company where patients can order diagnostic tests themselves with or without a health care professional. These diagnostic services are Directlab, where users can order web-based diagnostic test packages independent of a health care professional, and Homelab, where patients in the digital environment of their GP can order diagnostic test packages. In regular daily practice in the Netherlands, the patient needs to ask for a consultation with the GP (on the phone or in person) and ask for a diagnostic test for STIs. In this situation, the GP performs triage to identify whether it is necessary to conduct an STI test.

Data Analysis

To determine the diagnostic test advice of the digital triage tool, the characteristics of each vignette were entered into it. The ensuing advice was compared with the test advice of the GPs per vignette. To learn which factors influenced the GPs' decision-making process, the combination of the think-aloud process, vignettes, and semistructured interviews was used as a triangulation method to obtain a complete range of data to result in a strong conclusion [12,20]. All interviews were transcribed (intelligent) verbatim. When the transcripts were completed and uploaded to ATLAS.ti (version 22; ATLAS.ti Scientific Software Development GmbH), the audio recordings were deleted. In total, 2 authors (Fleur Rekveld and KS) conducted the qualitative data analysis according to the principles of thematic analysis. Fleur Rekveld and KS developed a preliminary coding scheme based on the coded data from the first 8 participants. The final coding scheme emerged after all the coding was performed by the 2 authors independently. The codes were grouped into themes and subthemes.

Results

Characteristics of the Study Population

Data saturation was reached after 10 interviews. The characteristics of the participants are presented in Table 1. Their ages ranged from 32 to 70 years, with a mean of 48.30 (SD 11.88) years. The number of men and women was almost equal (6/10, 60% and 4/10, 40%, respectively). Of the 10 GPs, 1 (10%) was retired, 3 (30%) were working part time as GPs, and 6 (60%) were working full time.

Table 1. Characteristics of the participants.

Participant	Age (y)	Gender	Employment status
1	32	Woman	Part time
2	55	Man	Full time
3	38	Man	Part time
4	59	Man	Full time
5	70	Man	Retired
6	53	Man	Full time
7	55	Woman	Full time
8	43	Man	Full time
9	38	Woman	Part time
10	40	Woman	Full time

Testing Advice of Digital Triage Tool Versus GPs

Table 2 shows, for each vignette, whether the digital tool would advise conducting an STI test and what each participant would advise to do. For 50% (3/6) of the vignettes (ie, numbers 1, 4, and 5), the digital triage tool's advice aligned with all participants' advice. For all 3 vignettes, the advice was to conduct a diagnostic test for STIs. For those 3 vignettes, the patients' risk factors were sufficiently clear for the participants to advise to conduct a test.

In vignette 1, the most important decision-making factor was the patient's age; young age combined with women was an important factor influencing the participants' test advice as having an STI could make this woman infertile. Participant 7 answered the following:

> *I would test her, always with women of her age who are sexually active.*

In addition, unsafe sex was an important factor in the decision to test.

XSL•FO
RenderX

For vignette 4, the main factor in advising to test was the "men having sex with men" risk factor. Participant 5 answered the following:

> *It is male-male contact, and in addition, there are changes in sexual contacts so that he can do an STI test.*

For vignette 5, all participants would advise conducting an STI test as well. Furthermore, 80% (8/10) mentioned that they would also conduct cervical cancer diagnostic tests because of the symptom of contact bleeding. Participant 9 mentioned the following:

> *In the case of contact bleeding, more research than only an STI is needed. It could be Chlamydia, but a smear test is needed to exclude cervical cancer.*

For the other 50% (3/6) of the vignettes, not all participants gave the same advice as each other or as the digital triage tool. For vignette 2, a total of 60% (6/10) of the participants agreed with the advice of the digital tool, and for vignettes 3 and 6, the proportions were 70% (7/10) and 80% (8/10), respectively. It is important to mention that the initial answer of the participants is presented in Table 2. It could be the case that participants answered "no" to advising an STI test for the patient initially. However, the participants mentioned that they would advise conducting an STI test after excluding other diseases. In addition, sometimes, the participants wanted more information about the patient's situation before advising to conduct an STI test.

For vignette 2, most participants wanted to know more about the patient's case before giving the advice to test for an STI. In addition, they wanted to conduct a physical examination or other tests, such as a test to exclude urinary infection, as the patient's symptoms seemed not totally compliant with those of an STI. Participant 2 said the following:

> *I would like to know a little more; why does he think he has an STI? Does he have other contacts next to his current relationship or an open relationship? Has he heard anything from his wife?*

Participant 4 answered the following:

> *I would check his urine.*

Participants answered that the symptoms and risk factors were too unclear to advise an STI test. A minority of the participants would test for an STI to exclude it or to satisfy the patient's request. Participant 2 answered the following:

> *He asked for an STI test so I would do one.*

The participants mentioned that, sometimes, a patient does not have an apparent reason for wanting to take an STI test or the patient has no symptoms that fit with those of an STI. However, sometimes patients do not want to discuss this in detail, and participants found it important to allow for testing at a low threshold if patients asked for it themselves. Participant 9 mentioned the following:

> *Maybe he (or his wife) is cheating, and they do not want to tell you that directly...It is always the question if the patient is honest with you, so I would test at a low threshold after I did a urine infection test, and then I think he would accept that.*

Table 2. Advice of the digital tool and the participants to test for a sexually transmitted infection.

	Digital triage tool	P[a]1	P2	P3	P4	P5	P6	P7	P8	P9	P10	Agreement, n (%)[b]
Vignette 1	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	10 (100)
Vignette 2	No	No	Yes	No	No	Yes	No	No	Yes	No	Yes	6 (60)
Vignette 3	Later	Later	Later	Later	Later	No	Yes	Yes	Later	Later	Later	7 (70)
Vignette 4	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	10 (100)
Vignette 5	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	10 (100)
Vignette 6	Yes	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes	No	Yes	8 (80)

[a]P: participant.

[b]Percentage of participants who agreed with the advice of the digital triage tool.

For vignette 3, most participants (7/10, 70%) answered that the patient could take an STI diagnostic test but at a later time. At this time, it was too early to detect an STI. A total of 20% (2/10) of the participants also mentioned that they would talk to the patient about her contraception and provide education about safe sex. Participant 2 said the following:

> *She had unsafe sex, so I would do two things. Maybe check if she uses birth control, and I would tell her that she can do an STI test after two weeks.*

Vignette 6 involved a patient from an endemic area. In total, 25% (2/8) of the participants who agreed with the advice of the digital tool mentioned the endemic area as a reason for testing. Participant 10 mentioned the following:

> *I would ask her some more questions; however, she is from Surinam, a risk area. So I would test her at a low threshold, especially for a serological test.*

The other 62% (6/8) of the participants mentioned low-threshold testing because of the patient's symptoms. Most participants (6/10, 60%) mentioned that they would check for a urinary infection, some before conducting an STI test and others in addition to it. Participant 1 mentioned the following:

> *I would check her urine first to ensure she has no urinary infection.*

XSL•FO
RenderX

It is important to note that almost all participants mentioned that, if a patient requested an STI test, they would meet the request. They also mentioned that, in some cases, they would also give patients more information about safe sex or conduct a physical examination. The decision to do so often depended on age or other risk factors such as contact bleeding. Especially in the case of younger patients, GPs educated them about safe sex and birth control. However, this information provision was not part of their decision-making process but rather of their consultation.

Extra Factors That Influenced the Decision of the GPs

There were several factors that the participants considered in their decision that were not included in the digital triage tool. The most important additional patient-related factors were anxiety about infection, the wishes of the patient, and age. Among all participants (10/10, 100%), the patient's anxiety was an additional reason for referring them to an STI test. The participants reasoned that a request for an STI test is not made easily and that there may be an unknown reason behind it. In their opinion, when patients experience fear-related stress, it might harm their health. Participant 10 mentioned the following:

> *Sometimes you feel that there is more than they want to say, and then you decide to test at a low threshold.*

Age played a role in the decision-making process of the GPs. This was especially the case in vignettes 1 and 3. The GPs mentioned that checking for STIs was important at a fertile age, especially for women. In the Dutch medical guidelines, it is noted that, below the age of 25 years, there needs to be a low threshold for STI testing even if patients report no complaints. Participant 6 answered the following in the interview about vignette 3:

> *Especially in younger patients, you want to know what they know about sex and the transmission of STIs.*

In 2 vignettes, the GPs felt the need to ask additional questions or conduct a physical examination. The digital triage tool only provides advice on an STI test. However, the symptoms may also indicate a urinary tract infection or a stage of cervical cancer. These tests are not advised via the digital tool but were advised by the participants in this study for those 2 vignettes.

One GP also considered who had to pay for the test and whether it was affordable. Participant 3 mentioned the role of the payer or possible reimbursement in the decision. He answered the following about vignette 6:

> *If she wants to pay for a test and she wants to do a test...Then, she can do a test.*

In summary, it can be generally said that GPs in this study paid extra attention to patient-related factors such as fear of infection, desire to undergo the test, and young age when deciding whether to request an STI test.

Discussion

Principal Findings

In this study, we tried to identify whether the advice of a digital triage tool based on medical guidelines aligned with GPs' medical practice. The results showed that other factors, which are not part of the guidelines, played a role in the GPs' decision-making process when determining whether to advise an STI test for a patient. The most important additional patient-related factors were the patient's anxiety, wishes, and age. The GPs also considered who had to pay for the test and whether it was affordable. Finally, the GPs were willing in some vignettes to ask additional questions or conduct a physical examination. The most notable factors are discussed in this section and compared with the literature.

In line with other research, the GPs' decision to test depends sometimes on the anxiety and wishes of the patient [7]; these factors were not included in the studied digital triage tool. This additional aspect aligns with the research by Hajjaj et al [5,7]. In addition, our results align with those of a study that researched the barriers to following guidelines among GPs [6] that showed that the patient's preferences were considered more important than following guidelines.

The interviews showed that the age of the patients was an important factor that influenced the GPs' advice. Specifically, younger age was an important reason to advise an STI test because of the risk of infertility and the sexual activity in this group. Age was not included as a factor in the digital triage tool. As STIs mainly occur under the age of 30 years, it is not surprising that GPs tend to advise testing more for patients in this age group [21].

From the literature, it was found that the factor "knowing the patient" influences the decision-making process of GPs [22]. Accumulated knowledge about the patient influences the context and interpretation of the conversation between the patient and the health care professional, especially in the case of psychosocial or unspecific problems such as fatigue. However, in this study, knowing the patient was not a factor that was considered in the vignettes. For this reason, the decisions that the GPs made in this study could be different in real life as they might know the patients.

In addition to patient-related factors (eg, the wishes of the patient), GP-related factors also influenced the decision-making process. The extent to which GPs were open to discussion with patients about why they wanted an STI test or to which GPs were willing to address patients' concerns influenced the decision. In addition, based on the findings of this study, it seems that the GPs expressed a preference for obtaining a complete set of information before deciding. For example, some GPs wanted to have more information about the situation of the patients and their partners. In some cases, GPs wanted to conduct a physical examination or other diagnostic tests (eg, urinary infection) to exclude other diseases. The digital triage tool is strictly bound to the guidelines set up without paying attention to, for example, the anxiety of the patient or the need for additional information. Other guidelines have been developed for possible symptoms of urinary tract infection or cervical problems, which have not yet been combined on the internet.

The advice of the digital triage tool is straightforward and always in line with a strict algorithm. In this study, GPs were found to recommend a diagnostic test for STIs more often than

XSL•FO
RenderX

the digital tool. In the Netherlands, a study showed that unnecessary diagnostics (overdiagnostics) are a common problem among Dutch GPs; slightly more than half of the participating GPs indicated that patients could submit a complaint for not requesting an examination that was indicated and that this played a role to some or a significant extent in the request for diagnostic testing [23].

Our study did not investigate whether the digital tool can prevent overdiagnostics, but we assume that it can be a powerful decision support tool for daily general practice, just as tools for pharmacotherapy are already in use. More research is needed to confirm this.

Another possible reason why GPs are more inclined to test seems to be that it could save them time [24]. For example, if a patient has vague symptoms, it would be easy to request some tests first without having a thorough conversation. Another possible reason specifically for low-threshold STI testing could be feelings of embarrassment to ask about sexual behavior [25]. Recently, a Dutch center for sexual health found that talking about sexual behavior is not done as often as it should by health care professionals [26]. This could be seen as an additional justification for supporting GPs with digital tools for STI testing.

This study does not suggest that digital triage is the holy grail to prevent overdiagnostics or that it is *the* solution to lower the work pressure of GPs. However, this vignette study confirms that GPs have a more holistic approach to their patients compared with a digital triage tool. A digital triage tool primarily relies on specific responses to predefined questions, whereas a GP can consider more factors such as social factors, lifestyle, and personal context. On the one hand, the comprehensive perspective of GPs might result in a higher frequency of diagnostics when compared with a digital triage tool. This is due to the GPs considering additional factors. Given the high workload and time constraints of GPs, the investigated digital tool can play a helpful role in daily decision-making. In contrast, this holistic approach by GPs could potentially lead to fewer diagnostics. Given their deep understanding of the patients' condition, GPs are better positioned to assess the necessity of tests.

This study has several limitations. It could be that social desirability influenced the GPs' answers on the vignettes and interviews. Potentially, the advice of the GPs was more in line with the guidelines compared with that in their daily practice as they were aware of the fact that they were part of research on this topic [12]. It is also worth mentioning that there could be a disparity between what people think they would do in a particular situation and their actual behavior [27]. In addition, this study is not generalizable to the entire field of diagnostics at general practices because of its focus on STI testing. As a starting point, this study identified factors that influenced the decision-making process of GPs for STI testing. In future research, we recommend investigating digital tools and the decision-making process of GPs for other common diagnostic tests.

A strength of this study is the combination of the vignette method, the think-aloud process, and the semistructured interviews, which aimed to obtain a complete range of data on the topic (triangulation). Although no actual patients were included in this study, we aimed to make the vignettes as valid as possible by developing and testing them with GPs. In addition, providing the same vignettes to different GPs made it easier to compare patients within different general practices instead of comparing real-life patients with different complaints and characteristics. Currently, we are working on a real-life study in which patients in the waiting room of a GP's office complete digital triage for STI testing (the result of the digital triage tool is not shown to the patient), after which they go on to have their planned consultation with the GP. At this consultation, the GP will also advise whether to test for an STI; the advice of the digital tool and of the GP will be compared. We expect more detailed and practical information to further refine this working method using a digital tool.

A qualitative study in which GPs were interviewed about their general attitude toward the use of digital tools by patients in their practice showed that GPs' attitudes toward digital STI diagnostic services were positive, and they acknowledged that the use of eHealth in their practice could result in a more efficient workflow [28].

It will be interesting to further investigate whether GPs are also willing to use digital triage tools as a standard gateway for their practice for some diagnostic tests. When a digital triage tool is implemented and integrated into the care pathway, it is important to investigate what users think of this integration and whether they are satisfied with this change in their way of working. For future research, it could be beneficial to make a comparison of the experiences of patients with a digital triage tool, triage at the GP's office, and a mix. Notably, recent studies on digital chatbots for medical questions have shown that patients perceived the chatbot's responses to be superior to those provided by GPs [29]. For future applications, it is essential to consider patients' eHealth literacy before using a digital triage tool as the primary tool in daily general practice [30,31]; hybrid care might be a solution to address all types of patients. Finally, it is important to realize that the tool in the care pathway needs to stay up-to-date and needs to be changed when the medical guidelines are updated [32]. This study showed that (holistic) factors that are not part of the digital triage tool affect GPs' decision-making. This is an interesting topic for future research as digital tools and artificial intelligence are increasingly being used in health care. Nowadays, GPs use digital medication prescription tools to support their decision-making, which could help with handwriting errors but also with poor treatment decisions [33]. Another example is an artificial intelligence system that could help GPs decide on the early detection of skin cancer [34,35]. Digital technologies such as these should be researched carefully to see what the impact and consequences are for both GPs and patients.

Conclusions

This study shows that, in some cases, patients receive different advice to undergo an STI test from a digital tool and from a GP. Other factors that are not part of medical guidelines play a role in the GPs' decision-making process when deciding whether to request an STI test. The most important additional patient-related factors were the patient's anxiety, wishes, and age. One GP also

XSL•FO
RenderX

considered who had to pay for the test and whether it was affordable. Finally, some GPs expressed a desire to ask additional questions or conduct a physical examination in certain vignettes. In comparison, the digital triage tool adhered more closely to the medical guidelines, with GPs being more inclined than the digital tool to recommend an STI test for the same patient case. Alignment between the digital tool and GP advice only occurred when the risk factors for STI testing were unequivocally evident. This confirms that GPs decide from a holistic perspective. On the basis of these initial findings, we cautiously posit that a digital triage tool for STI testing can potentially support GPs and may even serve as a substitute for in-person consultations in the future. However, it is imperative to conduct further research to establish safe and effective methods for implementing such a transition.

These conclusions should be approached carefully, recognizing that this study represents an initial exploration and that additional research is required to substantiate and refine these findings.

Acknowledgments

The authors would like to thank all general practitioners who participated in this study. They appreciate the time they invested in this research. In addition, the authors want to thank Fleur Rekveld for her contribution to this paper. Funding was provided by Unilabs. Unilabs played no role in the research design or data collection, analysis, or interpretation.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Translated vignettes from Dutch to English.
[DOCX File , 13 KB - humanfactors_v11i1e49221_app1.docx]

Multimedia Appendix 2
Semistructured interview protocol.
[DOCX File , 16 KB - humanfactors_v11i1e49221_app2.docx]

References

1. Eysenbach G. What is e-health? J Med Internet Res 2001 Jun 18;3(2):E20 [FREE Full text] [doi: 10.2196/jmir.3.2.e20] [Medline: 11720962]
2. Fang ML, Walker M, Wong KL, Sixsmith J, Remund L, Sixsmith A. Future of digital health and community care: exploring intended positive impacts and unintended negative consequences of COVID-19. Healthc Manage Forum 2022 Sep 01;35(5):279-285 [FREE Full text] [doi: 10.1177/08404704221107362] [Medline: 35775162]
3. Gottliebsen K, Petersson G. Limited evidence of benefits of patient operated intelligent primary care triage tools: findings of a literature review. BMJ Health Care Inform 2020 May 07;27(1):e100114 [FREE Full text] [doi: 10.1136/bmjhci-2019-100114] [Medline: 32385041]
4. Schmieding ML, Kopka M, Schmidt K, Schulz-Niethammer S, Balzer F, Feufel MA. Triage accuracy of symptom checker apps: 5-year follow-up evaluation. J Med Internet Res 2022 May 10;24(5):e31810 [FREE Full text] [doi: 10.2196/31810] [Medline: 35536633]
5. Hajjaj FM, Salek MS, Basra MK, Finlay AY. Nonclinical influences, beyond diagnosis and severity, on clinical decision making in dermatology: understanding the gap between guidelines and practice. Br J Dermatol 2010 Oct;163(4):789-799. [doi: 10.1111/j.1365-2133.2010.09868.x] [Medline: 20854402]
6. Carlsen B, Glenton C, Pope C. Thou shalt versus thou shalt not: a meta-synthesis of GPs' attitudes to clinical practice guidelines. Br J Gen Pract 2007 Dec 01;57(545):971-978 [FREE Full text] [doi: 10.3399/096016407782604820] [Medline: 18252073]
7. Hajjaj FM, Salek MS, Basra MK, Finlay AY. Non-clinical influences on clinical decision-making: a major challenge to evidence-based practice. J R Soc Med 2010 May 30;103(5):178-187 [FREE Full text] [doi: 10.1258/jrsm.2010.100104] [Medline: 20436026]
8. O'Donnell CA. Variation in GP referral rates: what can we learn from the literature? Fam Pract 2000 Dec;17(6):462-471. [doi: 10.1093/fampra/17.6.462] [Medline: 11120716]
9. Duddy C, Wong G. Efficiency over thoroughness in laboratory testing decision making in primary care: findings from a realist review. BJGP Open 2021 Apr;5(2):bjgpopen20X101146 [FREE Full text] [doi: 10.3399/bjgpopen20X101146] [Medline: 33293413]
10. Versluis A, Schnoor K, Chavannes NH, Talboom-Kamp EP. Direct access for patients to diagnostic testing and results using eHealth: systematic review on ehealth and diagnostics. J Med Internet Res 2022 Jan 12;24(1):e29303 [FREE Full text] [doi: 10.2196/29303] [Medline: 35019848]

XSL•FO
RenderX

11. Balfe M, Brugha R. Disclosure of STI testing activities by young adults: the influence of emotions and social networks. Sociol Health Illn 2010 Nov 11;32(7):1041-1058 [FREE Full text] [doi: 10.1111/j.1467-9566.2010.01281.x] [Medline: 20937054]
12. Daemers DO, van Limbeek EB, Wijnen HA, Nieuwenhuijze MJ, de Vries RG. Factors influencing the clinical decision-making of midwives: a qualitative study. BMC Pregnancy Childbirth 2017 Oct 06;17(1):345 [FREE Full text] [doi: 10.1186/s12884-017-1511-5] [Medline: 28985725]
13. Hennink M, Kaiser BN. Sample sizes for saturation in qualitative research: a systematic review of empirical tests. Soc Sci Med 2022 Jan;292:114523 [FREE Full text] [doi: 10.1016/j.socscimed.2021.114523] [Medline: 34785096]
14. Fonteyn ME, Kuipers B, Grobe SJ. A description of think aloud method and protocol analysis. Qual Health Res 2016 Jul 01;3(4):430-441 [FREE Full text] [doi: 10.1177/104973239300300403]
15. Lundgrén-Laine H, Salanterä S. Think-aloud technique and protocol analysis in clinical decision-making research. Qual Health Res 2010 Apr 03;20(4):565-575. [doi: 10.1177/1049732309354278] [Medline: 19959822]
16. Atzmüller C, Steiner PM. Experimental vignette studies in survey research. Methodol 2010 Jan;6(3):128-138 [FREE Full text] [doi: 10.1027/1614-2241/a000014]
17. Van Bergen J, Dekker J, Boeke A, Kronenberg E, Van der Spruit R, Burgers J. NHG-Standaard Het soa-consult (eerste herziening). Huisarts Wet 2013;56(9):450-463. [doi: 10.1007/978-90-313-8279-8_88]
18. Home page. Saltro. URL: https://saltro.nl/ [accessed 2024-01-12]
19. Home page. National eHealth Living Lab. URL: https://nell.eu/ [accessed 2024-01-12]
20. Noble H, Heale R. Triangulation in research, with examples. Evid Based Nurs 2019 Jul 14;22(3):67-68. [doi: 10.1136/ebnurs-2019-103145] [Medline: 31201209]
21. Zorg V. SOA: Leeftijd & Geslacht. VZinfo. 2017. URL: https://www.vzinfo.nl/soa/leeftijd-en-geslacht/algemeen [accessed 2024-01-12]
22. Hjortdahl P. The influence of general practitioners' knowledge about their patients on the clinical decision-making process. Scand J Prim Health Care 1992 Dec 12;10(4):290-294. [doi: 10.3109/02813439209014076] [Medline: 1480869]
23. Wammes J, Verhoef L, Westert G, Assendelft P, Jeurissen P, Faber M. Onnodige zorg in de Nederlandse gezondheidszorg, gezien vanuit het perspectief van de huisarts. Celsus Academie voor Betaalbare Zorg. 2013. URL: https://betaalbaarheidvanzorg.nl/images/Publicaties/CELSUS_Werkdocument_Onnodige%20zorg%20vanuit%20perspectief%20huisarts_DEF.pdf [accessed 2024-01-12]
24. Vervloet M, Bomhoff M, Schellevis F, van Dijk L. Niet te veel en niet te weinig: de balans tussen nodige en onnodige zorg in de huisartsenpraktijk. NIVEL. 2015. URL: https://www.nivel.nl/sites/default/files/bestanden/Overzichtsstudie-niet-te-veel-en-niet-te-weinig.pdf [accessed 2024-01-12]
25. Hinchliff S, Gott M, Galena E. GPs' perceptions of the gender-related barriers to discussing sexual health in consultations--a qualitative study. Eur J Gen Pract 2004 Jun 11;10(2):56-60. [doi: 10.3109/13814780409094233] [Medline: 15232525]
26. Zuil W. Rutgers: seksualiteit onvoldoende besproken binnen de zorg. Skipr. URL: https://www.skipr.nl/nieuws/rutgers-seksualiteit-onvoldoende-besproken-binnen-de-zorg/?daily=1&utm_medium=email&utm_source=20230425+skipr+daily+nieuwsbrief-daily&utm_campaign=NB_SKIPR&tid=TIDP1993360X0AD5FC84E9664D39893DF4F2104FF76CYI4 [accessed 2024-01-12]
27. Erfanian F, Roudsari RL, Heydari A, Bahmani MN. A narrative on using vignettes: its advantages and drawbacks. J Midwifery Reproductive Health 2020;8(2):2134-2145 [FREE Full text] [doi: 10.22038/JMRH.2020.41650.1472]
28. de Wilt T, Versluis A, Goedhart A, Talboom-Kamp E, van Delft S. General practitioners attitude towards the use of eHealth and online testing in primary care. Clin eHealth 2020;3:16-22 [FREE Full text] [doi: 10.1016/j.ceh.2020.02.002]
29. Ayers JW, Poliak A, Dredze M, Leas EC, Zhu Z, Kelley JB, et al. Comparing physician and artificial intelligence chatbot responses to patient questions posted to a public social media forum. JAMA Intern Med 2023 Jun 01;183(6):589-596. [doi: 10.1001/jamainternmed.2023.1838] [Medline: 37115527]
30. van der Kleij RM, Kasteleyn MJ, Meijer E, Bonten TN, Houwink EJ, Teichert M, et al. SERIES: eHealth in primary care. Part 1: concepts, conditions and challenges. Eur J Gen Pract 2019 Oct 10;25(4):179-189 [FREE Full text] [doi: 10.1080/13814788.2019.1658190] [Medline: 31597502]
31. Heijsters F, van Loon G, Santema J, Mullender M, Bouman M, de Bruijne M, et al. A usability evaluation of the perceived user friendliness, accessibility, and inclusiveness of a personalized digital care pathway tool. Int J Med Inform 2023 Jul;175:105070 [FREE Full text] [doi: 10.1016/j.ijmedinf.2023.105070] [Medline: 37121138]
32. Stahl L, Spatz M. Quality assurance in eHealth for consumers. J Consum Health Internet 2003 Jan;7(1):33-42. [doi: 10.1300/j381v07n01_03]
33. Armando LG, Miglio G, de Cosmo P, Cena C. Clinical decision support systems to improve drug prescription and therapy optimisation in clinical practice: a scoping review. BMJ Health Care Inform 2023 May 02;30(1):e100683 [FREE Full text] [doi: 10.1136/bmjhci-2022-100683] [Medline: 37130626]
34. Micocci M, Borsci S, Thakerar V, Walne S, Manshadi Y, Edridge F, et al. Attitudes towards trusting artificial intelligence insights and factors to prevent the passive adherence of GPs: a pilot study. J Clin Med 2021 Jul 14;10(14):3101 [FREE Full text] [doi: 10.3390/jcm10143101] [Medline: 34300267]

XSL•FO
RenderX

35. Sangers TE, Wakkee M, Moolenburgh FJ, Nijsten T, Lugtenberg M. Towards successful implementation of artificial intelligence in skin cancer care: a qualitative study exploring the views of dermatologists and general practitioners. Arch Dermatol Res 2023 Jul 07;315(5):1187-1195 [FREE Full text] [doi: 10.1007/s00403-022-02492-3] [Medline: 36477587]

Abbreviations

GP: general practitioner
STI: sexually transmitted infection

Edited by A Kushniruk; submitted 30.05.23; peer-reviewed by S Hinder; comments to author 30.06.23; revised version received 05.07.23; accepted 20.11.23; published 22.01.24.

Please cite as:
Schnoor K, Versluis A, Chavannes NH, Talboom-Kamp EPWA
Digital Triage Tools for Sexually Transmitted Infection Testing Compared With General Practitioners' Advice: Vignette-Based Qualitative Study With Interviews Among General Practitioners
JMIR Hum Factors 2024;11:e49221
URL: https://humanfactors.jmir.org/2024/1/e49221
doi:10.2196/49221
PMID:38252474

©Kyma Schnoor, Anke Versluis, Niels H Chavannes, Esther P W A Talboom-Kamp. Originally published in JMIR Human Factors (https://humanfactors.jmir.org), 22.01.2024. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on https://humanfactors.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Evaluating Users' Experiences of a Child Multimodal Wearable Device: Mixed Methods Approach

Nancy L McElwain[1,2], DPhil; Meghan C Fisher[1], DPhil; Camille Nebeker[3], MS, EdD; Jordan M Bodway[1], BSc; Bashima Islam[4], DPhil; Mark Hasegawa-Johnson[2,5], DPhil

[1]Department of Human Development and Family Studies, University of Illinois Urbana-Champaign, Urbana, IL, United States
[2]Beckman Institute for Advanced Science & Technology, University of Illinois Urbana-Champaign, Urbana, IL, United States
[3]Herbert Wertheim School of Public Health and Human Longevity Science, University of California, San Diego, La Jolla, CA, United States
[4]Department of Electrical and Computer Engineering, Worcester Polytechnic Institute, Worcester, MA, United States
[5]Department of Electrical and Computer Engineering, University of Illinois Urbana-Champaign, Urbana, IL, United States

Corresponding Author:
Nancy L McElwain, DPhil
Department of Human Development and Family Studies
University of Illinois Urbana-Champaign
2004 Doris Christopher Hall
904 West Nevada Street
Urbana, IL, 61801
United States
Phone: 1 217 244 7168
Email: mcelwn@illinois.edu

Abstract

Background: Wearable devices permit the continuous, unobtrusive collection of data from children in their natural environments and can transform our understanding of child development. Although the use of wearable devices has begun to emerge in research involving children, few studies have considered families' experiences and perspectives of participating in research of this kind.

Objective: Through a mixed methods approach, we assessed parents' and children's experiences of using a new wearable device in the home environment. The wearable device was designed specifically for use with infants and young children, and it integrates audio, electrocardiogram, and motion sensors.

Methods: In study 1, semistructured phone interviews were conducted with 42 parents of children aged 1 month to 9.5 years who completed 2 day-long recordings using the device, which the children wore on a specially designed shirt. In study 2, a total of 110 parents of children aged 2 months to 5.5 years responded to a questionnaire assessing their experience of completing 3 day-long device recordings in the home. Guided by the Digital Health Checklist, we assessed parental responses from both studies in relation to the following three key domains: (1) access and usability, (2) privacy, and (3) risks and benefits.

Results: In study 1, most parents viewed the device as easy to use and safe and remote visits as convenient. Parents' views on privacy related to the audio recordings were more varied. The use of machine learning algorithms (vs human annotators) in the analysis of the audio data, the ability to stop recordings at any time, and the view that the recordings reflected ordinary family life were some reasons cited by parents who expressed minimal, if any, privacy concerns. Varied risks and benefits were also reported, including perceived child comfort or discomfort, the need to adjust routines to accommodate the study, the understanding gained from the study procedures, and the parent's and child's enjoyment of study participation. In study 2, parents' ratings on 5 close-ended items yielded a similar pattern of findings. Compared with a "neutral" rating, parents were significantly more likely to agree that (1) device instructions were helpful and clear (t_{109}=−45.98; P<.001), (2) they felt comfortable putting the device on their child (t_{109}=−22.22; P<.001), and (3) they felt their child was safe while wearing the device (t_{109}=−34.48; P<.001). They were also less likely to worry about the audio recordings gathered by the device (t_{108}=6.14; P<.001), whereas parents' rating of the burden of the study procedures did not differ significantly from a "neutral" rating (t_{109}=−0.16; P=.87).

Conclusions: On the basis of parents' feedback, several concrete changes can be implemented to improve this new wearable platform and, ultimately, parents' and children's experiences of using child wearable devices in the home setting.

(JMIR Hum Factors 2024;11:e49316) doi:10.2196/49316

XSL•FO
RenderX

KEYWORDS

wearable devices; multimodal sensing; user experience; usability; privacy; children; mobile phone

Introduction

Background

Advances in pervasive sensing, internet of medical things, and digital health strategies more broadly [1-6] have rapidly accelerated over the past decade. Although digital health research among adults and adolescents has predominantly used smartphones [7-9], parallel work with infants and children tends to use wearable devices [10], including motion sensors to detect body posture and physical activity [11], audio recorders to assess language environment and development [12,13], heart rate sensors to assess psychophysiology [14], and head-mounted cameras to capture infants' visual perspective of the physical and social environment [15]. Such wearable technology, especially when paired with machine learning algorithms, permits the automated detection of children's behavioral and physiological states, as well as caregivers' responses, and has the potential to transform the field of child development through the collection of big data in real-world environments [16].

At the same time, the use of wearable devices among infants and young children in home environments raises unique ethical, legal, and social implications and logistical challenges. As such, careful attention to the perspectives and experiences of end users of such technology, in this case, parents and their children, is required. In this study, we assessed parents' perceptions of and experiences with a novel wearable device, LittleBeats, developed specifically for use with infants and young children. Little Beats, which is not Food and Drug Administration approved and used only for research purposes, integrates a microphone, a 3-lead electrocardiogram (ECG) sensor, and an inertial motion sensor to synchronously collect information about infant vocalizations, cardiac physiology (heart rate and respiratory sinus arrhythmia), and motion (eg, physical activity level, position, and discrete movements). The electronics are housed in a 3D-printed case (55×57×13 mm), which is placed on a specially designed shirt that the child wears. Data can be collected throughout the day at home, without the researchers present. In prior papers, we reported on machine learning algorithms used to detect and classify child and parent vocalizations using audio data from the LittleBeats device [17] and child sleep states using all 3 sensor modalities [18]. We also conducted technical validation studies to assess the signal quality of each sensor modality in relation to established laboratory protocols and gold-standard equipment [19]. Complementing these prior reports, we focus here on the critical issue of "user experience" among families and their children aged 1 month to 9.5 years. Using semistructured interviews and parent questionnaires to assess parents' experiences and perceptions, our mixed methods investigation examined usability, privacy, and perceived risks and benefits.

The "Digital Health Checklist" for Use in Child Development Research

The proliferation of digital health technologies has spurred a parallel examination of ethical practices and related decision-making processes around the use of such technologies with human participants. To evaluate the LittleBeats platform, we used the Digital Health Checklist developed by Nebeker et al [20,21]; it is grounded in the ethical principles of the *Belmont Report* [22], which speaks to beneficence, respect for persons (or autonomy), and justice, and the *Menlo Report* [23], which added the principle of respect for law and public trust. These principles form the foundation of a 4-domain framework that includes privacy, access and usability, data management (eg, collection, storage, interoperability, and sharing), and assessment of risks and benefits (Figure 1).

To date, the research and development of the Digital Health Checklist has been applied to digital health protocols in adult samples, including for use in cardiovascular disease prevention [24]; studies of human emotion [25]; and improvement of informed consent communications [26]. The current investigation extended the use of the Digital Health Checklist to research involving parents of infants and children. In doing so, we integrated ethical considerations specific to research with children [27]. Specifically, children are a heterogeneous group, and the potential benefits and risks to child participants need to be understood within the context of the child's age and related physical, cognitive, and socioemotional abilities.

For instance, infants and toddlers may be more susceptible to risks related to emotionally stressful procedures because their coping abilities are less well developed and depend, in part, on support from caregivers. By contrast, older children may be better able to regulate emotions and exert their autonomy, although they might be at an increased risk in other domains. For instance, owing to their growing self-awareness and other awareness, preschool- and school-aged children may be increasingly susceptible to experiencing shame and embarrassment, heightened concerns about privacy, and other related risks to the child's self-concept.

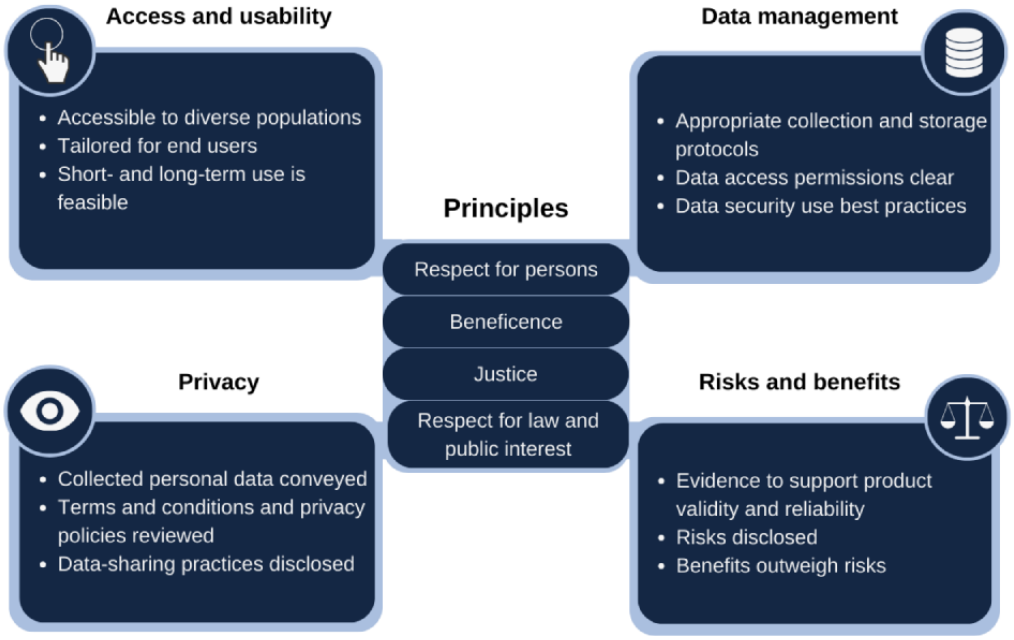
With developmental differences in risk assessment in mind, we assessed LittleBeats user experience among children representing a large age range (infancy through middle childhood). Although we did not interview children about their study experiences, we considered the children's age in our analysis of parents' open-ended responses and parents' perspectives regarding how their children felt about and responded to the research procedures. Research with children requires parental consent and, depending on the child's age, the child's assent to affirm their willingness to participate in the research. The consent process related to the LittleBeats technology has been addressed in a prior report [26]; therefore, we did not consider issues related to the provision of parental consent before participating in this research. Instead, our focus here was on parents' perceptions of and reflections on their own and their children's experiences following the use of LittleBeats at home.

Although child development research incorporating the use of wearable devices is rapidly expanding [28-32], systematic assessment of parents' perspectives and experiences (or ethical

XSL•FO
RenderX

considerations more broadly) of such research has been sparse. A notable exception is a report by Levin et al [33], which outlines several key concerns parents may have about participating in research using wearable or remote sensing devices. These concerns focus on privacy expectations, particularly regarding audio or video data (considered "high fidelity data streams"), data management, and data use (eg, for scientific vs commercial purposes). Although we know of no study that assessed parents' perceptions and experiences of using wearable devices at home *after* data collection, Levin et al [33] provided valuable insights into parents' general willingness to participate in such research. Among a nationally representative sample of 210 parents (n=105, 50% mothers) with at least 1 child aged ≤5 years, 71.4% (n=150) of parents responding to hypothetical scenarios indicated at least some willingness to participate in studies involving motion or physiological sensors (low fidelity), whereas a significantly lower percentage of parents (n=99, 47.1%) endorsed willingness to participate in studies gathering audio recordings at home. It remains unknown whether the concerns expressed in the study by Levin et al [33], in which parents hypothetically considered participating in different types of remote sensing research, would also be voiced among parents who participated in research in which their children wore a wearable device with multiple sensor types (eg, motion, physiology, and audio).

Figure 1. Four-domain framework of the Digital Health Checklist for researchers. The Digital Health Checklist for researchers depicts the 4 ethical principles undergirding the 4 key domains of access and usability, privacy, risks and benefits, and data management. Source: this figure is published with permission and reflects an adaptation of the Digital Health Checklist Developed for Researchers (DHC-R) [34,35].



This Study

Guided by the domains of the Digital Health Checklist [20], we assessed parents' experiences with and perceptions of using LittleBeats at home using a mixed methods approach. In study 1, we conducted a qualitative (thematic) analysis of parental responses to a semistructured interview following the completion of 2 day-long LittleBeats recordings at home; children in this study were aged between 1 month and 9.5 years. In study 2, we collected data on parents' perspectives of using LittleBeats (again, following the completion of several day-long recordings at home) from a separate, larger sample. In this second study, we administered close-ended questionnaire items developed considering the qualitative themes identified in study 1. The parents in study 2 also had the opportunity to provide open-ended comments. In study 2, we narrowed our developmental focus to children aged 1 month to 5 years because our substantive interests focused on early childhood, and analytic tools are currently being developed for LittleBeats data collected among children aged ≤5 years.

Study 1

Methods

Participants

A total of 47 families with children aged 1 month to 9.5 years were recruited through web-based forums (eg, Facebook [Meta Platforms, Inc] parenting groups) and flyers distributed to local organizations (eg, libraries and day care centers) in a small Midwestern city. Because the larger study from which data were drawn included assessments of child stress physiology, families were excluded if their children had any known cardiac abnormalities. Of the 47 families that participated in the larger study, 42 (89%) completed the follow-up interview about their experience of using LittleBeats at home. Interviews were not completed with 5 (11%) families because of losing contact with them or because interview procedures were not finalized at the time of their study participation.

From these 42 families, 43 children (n=20, 47% female) participated. In 1 instance, 2 (5%) children (aged 13 and 71 mo) were from the same family. Children were aged 1.1 month to

XSL•FO
RenderX

9.5 years (mean 44.9, SD 38.36 mo) and represented 6 age groups: young infants (aged 1-5 mo; 7/43, 16%), older infants (aged 6-17 mo; 10/43, 23%), toddlers (aged 18-35 mo; 7/43, 16%), preschool-aged children (aged 36-59 mo; 6/43, 14%), early school-aged children (aged 5-7 y; 7/43, 16%), and school-aged children (aged 8-10 y; 6/43, 14%). Overall, 22 (51%) children were first born, 11 (26%) were second born, and 9 (21%) were third or later born. Mothers were aged, on average, 35.04 (SD 4.09) years, and fathers were aged, on average, 37.42 (SD 4.48) years. Across mothers and fathers, the highest level of education reported included a high-school degree (1/79, 1%), some college or 2-year degree (18/79, 23%), a bachelor's degree (22/79, 28%), or an advanced degree (38/79, 48%). Parents identified as Black (2/79, 3%), Asian (3/79, 4%), White non-Hispanic (70/79, 89%), Hispanic (2/79, 3%), or >1 race (2/79, 3%). These demographic data were missing for 2 (5%) of the 42 mothers and 3 (7%) of the 42 fathers. The mean family income was US $79,500 (SD US $25,000).

Ethical Considerations

This study was approved by the institutional review board at the University of Illinois Urbana-Champaign (protocol #21032).

Overview of LittleBeats Procedures

LittleBeats collects 3 streams of data (ECG, motion, and audio data) simultaneously while participants go about their everyday routines (Figure 2). Owing to COVID-19 protocols, all participant engagement was remote. LittleBeats kits (ie, LittleBeats device and shirt, ECG leads, disposable ECG electrodes, alcohol swabs to remove residue from electrodes, medical tape to secure wires on the child's chest, charging cable and block, and setup instruction cards) were either mailed or delivered by a research coordinator to the family's home. After receiving the kit, the mother and child met with the study coordinator through Zoom, a secure video web-conferencing platform. During this 40-minute Zoom visit, the study coordinator guided the mother through the LittleBeats setup (described in more detail subsequently), and the mother-child dyad participated in a series of tasks (video recorded for subsequent coding), including a baseline assessment of child stress physiology at rest and a mother-child play session.

For child participants aged <7 years, the mother-child dyads were also asked to complete a brief series of age-appropriate motion interaction tasks, such as the mother picking up her child (aged 1-4 mo), the mother and child (aged 11 mo) clapping together, or the mother and child (aged 6 y) playing "Simon Says." Toward the end of the Zoom visit, the study coordinator provided instructions for completing the LittleBeats home recordings. Families were asked to complete 2 day-long recordings (approximately 8 hours per day). All adults present at home during the recordings (eg, parents, grandparents, and babysitters) were required to provide consent to the LittleBeats recordings using a secure web-based form provided by the research team. If any nonconsenting adults were at home, parents were asked to turn off the device while these individuals were present. At the end of each day of recording, parents (usually mothers) completed a brief questionnaire about the day's recording (eg, recording start and stop times). To compensate the families for their time, parents were sent a US $100 e-gift card.

With regard to setting up LittleBeats, the research coordinator walked the mother through the following setup steps at the beginning of the Zoom visit: (1) threading a set of ECG lead wires (20 cm) through the back of the shirt pocket, (2) connecting ECG leads jack (2.5 mm) to the LittleBeats device, (3) turning the device on by sliding the switch to the "on" position (confirmation that the beginning of the recording is indicated by a red flashing light displayed on the device), (4) placing the device in a snug, specially designed shirt pocket, which is secured using 2 snaps, (5) snapping leads to 3 repositionable latex-free gel electrodes, (6) putting the LittleBeats shirt on the child, (7) cleaning the skin (where the electrodes will be placed) with an alcohol prep pad and then placing the electrodes on the child's skin, and (8) applying a small strip of 3M Micropore medical tape to each ECG wire approximately 5.1 cm below each ECG sticker to help secure the wires in place.

At the end of the Zoom visit, the research coordinator also walked the mother through how the LittleBeats device should be removed. The removal steps include (1) removing electrodes from the child's skin and using provided alcohol wipes, as needed, to remove residual gel from the electrodes; (2) unsnapping the electrodes from the ECG wires; (3) taking off the LittleBeats shirt; (4) removing the device from the shirt pocket; (5) sliding the slide switch to the "off" position; and (6) plugging the device into the provided charging cable (microUSB cable).

Figure 2. (A) LittleBeats device case; (B) LittleBeats supplies, including electrocardiogram leads, electrodes, charger, and shirt; and (C) an infant wearing LittleBeats at home.

XSL•FO
RenderX

LittleBeats Device Design and Study Implementation for End Users

LittleBeats was developed with parents and children (ie, end users) in mind. To provide a context for parents' interview responses about their study experiences, we noted several aspects of the device design and study implementation intended to proactively increase usability and decrease concerns about privacy. With respect to usability, we provided participants with clear, illustrated instructions in several formats (eg, hard copy and on the web). The device was also designed to be simple to use, with an on-off switch and a charging port, and we provided parents with the all the materials in the LittleBeats kit (refer to the *Overview of LittleBeats Procedures* section) that they would need to set up and use LittleBeats at home. For the child's comfort, the device is compact (55×57×14 mm) and lightweight (1.48 oz), with foam padding lining the inside of the shirt pocket in which the device is to be placed. The shirts are adorned with a variety of pocket designs (eg, hearts, animals, trucks, and dinosaurs) to appeal to toddlers and preschool-aged children, and, as part of the LittleBeats kit, families received 2 shirts with different designs. For older children, we provided more age-appropriate solid shirt pockets.

With respect to privacy, audio recordings provide high-fidelity information regarding participants' lives and require special considerations related to participant privacy, data confidentiality, and recording bystanders (for review, refer to the study by Cychosz et al [13]). Our approach to protecting participant privacy aligns with user-centered privacy protections recommended for mobile health research [36] and a "rights-based" approach adopted increasingly in the United States and used by the European Union (ie, General Data Protection Regulation), according to which individuals have the right to control their personal data, including but not limited to consent, erasure, secure data management practices, and transparency. For example, an important strategy to minimize privacy risks includes giving participants control over recordings [37,38]. In this vein, parents were told at several points during their participation (eg, consent process and consent form, verbally during the Zoom visit, and written instruction card) that they were free to turn off or pause the device at any time and that they could request that their recordings be partially or fully destroyed and not be used in the research. With respect to third-party individuals, parents were also instructed to use the device at home when only immediate family members or other consenting adults are present. Parents were informed that all data files were marked only by identification numbers, machine learning algorithms would be used to process the audio data, research personnel would listen to only snippets of the audio files as part of checks on algorithm development and accuracy, and research personnel were trained to protect participant privacy and would immediately cease listening to audio snippets in instances where personal information (discussion of medical, financial, or other personal issues) is being relayed. To minimize the risk of data being intercepted during transfer (ie, uploading data via wireless or Bluetooth networks), data were stored directly on a microSD card on the physical device, and files were configured in such a way that only study personnel could access the data in a human-readable format (eg, wav files for audio) using a data processing pipeline developed specifically for LittleBeats. Because LittleBeats is not a commercial device, simple modifications can be made to the device firmware (eg, "turning off" ≥1 of the sensors) to suit research goals (refer to the study by Islam et al [19] for details about technical specifications).

Parent Interview and Coding Procedures

Upon the completion of the LittleBeats recordings, parents (41 mothers and 1 father) completed a brief phone interview about their experiences of using LittleBeats in the home. To help minimize social desirability biases in parental responses, such as parents' reports of positive experiences with LittleBeats instructions received during the Zoom visit, these interviews were conducted by a second study coordinator who was not present during the Zoom visit. Guided by the dimensions outlined in the Digital Health Checklist [20], as well as special considerations related to research with children [27], our semistructured interview was designed for the purpose of this research to capture information about parents' experiences and perspectives regarding access, usability, privacy concerns, and risks and benefits with respect to the use of the LittleBeats device and the process of carrying out home recordings. Participants rarely provided information specific to the fourth domain of the Digital Health Checklist, data management, which encompasses how data are collected, stored, and shared and the extent to which the data are accessible to other systems or interoperability. Given the nature of the LittleBeats data (ie, they are not shared outside the research team, not accessible or integrated with other systems, and not transferred via a wireless or Bluetooth network that might be susceptible to security breaches), the data management theme is somewhat less relevant to LittleBeats than to health applications that might be accessed by multiple users (eg, patients, health care providers, and insurance providers). When parents expressed their views on the processes of data collection, storage, and security in the interviews, they almost exclusively focused on the audio recordings and privacy considerations. Therefore, we coded these responses under the privacy domain.

The interview included 11 open-ended questions, and the study coordinator conducting the interviews used standard probes to gain more insight into parents' experiences, perceptions, concerns, and questions (Multimedia Appendix 1). The interview questions allowed for feedback from all family members' perspectives (ie, the participating child, participating parents, and any other children or adults in the home). All parent interviews, conducted by the same study coordinator to ensure consistency, were audio recorded with the participant's permission. Interview recordings were manually transcribed, and identifiable information (eg, names and birth dates) and conversational placeholders (eg, "uh-huh") were omitted from the transcripts.

We used Taguette [39], an open-source web-based tool for coding textual qualitative data, to capture prevalent themes in our interview data and followed the 6-step approach to thematic analysis defined by Braun and Clarke [40]. At step 1, a review of the transcripts provided preliminary ideas for codes. At step 2, initial codes were generated based on the data from 5

XSL•FO
RenderX

interview transcripts of parents with children from different age groups. Through a series of team discussions, we developed an initial codebook focusing on areas that fell into the larger categories outlined in the Digital Health Checklist [20]. Three transcripts were then used for training purposes, and 3 researchers individually coded the transcripts. Discrepancies were discussed, and additional changes were made to the codebook. Upon the completion of the training, 1 researcher (who was not informed of the specific study objectives) coded all the transcripts using the refined codebook. Reliability was assessed by having the fourth author code 8 randomly chosen transcripts, and among the parent responses that both coders deemed codable, agreement was excellent (Cohen κ=0.967). At step 3, the research team met on a regular basis throughout the coding process to identify and discuss potential themes. At step 4 and after the completion of coding, final themes were reviewed by checking themes in relation to the entire data set to ensure an accurate representation of the data. At step 5, themes were refined and finalized by providing descriptive labels and definitions. At the final step, we organized the results based on the key domains of the Digital Health Checklist and created a summary table of themes with selected interview excerpts to illustrate the findings.

Results

Overview

Themes identified under the major categories of access and usability, privacy, and risks and benefits are summarized in the subsequent sections. Overall, similar themes were identified across developmental periods, although specific examples illustrating a given theme often differed depending on whether the parent reported on their infant, toddler, preschool-aged child, or school-aged child.

Access and Usability

According to the Digital Health Checklist, the domain of access and usability prompts researchers to consider whether the participant will be able to use the device as intended. This may involve evaluating whether the product has infrastructure requirements, such as internet access, as well as whether the device has been successfully used in the target population. In this study, usability refers to parents knowing how and being able to successfully use the LittleBeats device and materials (eg, ECG leads). Furthermore, usability encompasses families' experience of and ability to adhere to the study procedures more generally (ie, participant burden, eg, completing multiple day-long recordings), beyond the use of the device itself (Table 1).

A majority of parents expressed sentiments regarding their ability to easily operate the device (ie, turning the device on-off and charging the device). Some parents indicated feeling comfortable given their previous experience with comparable equipment, yet other parents with no such prior experience expressed similar views about the ease of use. Parents also commented that the instructions were helpful and appreciated having a variety of resources to refer to, if needed (eg, written instruction card, website, and study personnel contact). Aside from operating the device itself, parents had varying views on the materials needed to place the device on their child. Some parents noted that the design was well thought out and that setting up the electrodes was not complicated. However, other parents indicated some challenges with the materials, such as with threading the electrodes through the back of the shirt pocket.

Parents also expressed differing perspectives about the ease of setting up (and removing) the device. Although many parents felt comfortable placing LittleBeats on their children, some parents noted that gaining their children's cooperation was sometimes a challenge. For instance, some parents reported difficulty putting the device on their "wiggly, squiggly" infants. Other parents reported reluctance on the part of their toddlers or preschool-aged children, who could express their opinions and desires verbally. Typically, if challenges related to child cooperation were experienced, it was during the setup phase, and parents suggested that once their child was wearing LittleBeats, it was quickly forgotten. Parents expressed that the placement of the device on the upper anterior torso (ie, chest) may be disruptive to some activities, such as napping for a child who is a tummy sleeper. Relatedly, the device being concealed in the shirt pocket, with the ECG leads underneath the shirt, was viewed as a disadvantage by some parents who wanted to know whether the device was recording properly or whether there was a malfunction (eg, device turned off or ECG electrodes fell off).

With respect to participant burden, parents expressed a mix of perspectives. Many parents described day-long recordings (ie, >8 h/d) as feasible but challenging. However, parents noted factors that mitigated this challenge, such as the need to record for only a limited number of days spaced across multiple weeks, the ability to schedule their recordings when it worked for them, and the reduction in other competing activities due to the COVID-19 pandemic. In the same vein, parents expressed wanting more features to help them fulfill project expectations. Currently, the device provides no information to the user beyond an indicator light showing that the device is powered on. Parents found it difficult to know how long they had recorded for or how much battery charge was left when using the device.

In addition, many parents described the project as convenient, indicating that the remote data collection procedures were appealing. Being able to collect data at home, on their family's own schedule, made it relatively easy to participate. Parents were not burdened by the need to travel to a research laboratory, and they could set up the device and start recording when it fit their schedule. Concerns about being able to keep the device on securely or ensure that the device was collecting data were voiced by some parents of older and more active children (eg, increased unsupervised time and gel adhesive weakening owing to perspiration). Other parents expressed their worry that their children would damage the device during data collection.

XSL•FO
RenderX

Table 1. Themes, subthemes, and example excerpts related to the access and usability of the LittleBeats device and study procedures (study 1).

Themes and subthemes	Example excerpts
Operating the device	• "Everything was pretty easy. It was easy to charge, it was easy to you know put the stickers on and attach, and like I said I don't think she really felt like it was on. The first day after she asked after an hour 'how long have I had it on' I was like 'why is it uncomfortable' she was like 'no I was just wondering' and I was like 'oh okay.' I don't think she even realized she had it on half the time." (Parent of a school-aged child)
Instructions	• "They [the instructions] were very clear. I mean they made it so that I felt confident putting it on her and doing what I was supposed to do." (Parent of an older infant)
Support materials (specially designed t-shirt, wires, and electrodes)	
Ease of use	• "I think the t-shirt definitely made it easier to use. That was a nice little set up, and it made it, you know, stay in place and like see where it [the device] needed to be for it to be hooked up and stay in place...And then even with the hole on the inside [of the shirt] to make it easy to get all the cords. That was really a unique design tool but effective." (Parent of a school-aged child)
Challenges	• "It was a little hard getting the black metal piece through the back of the shirt. Like I needed that hole to be a little bigger. So, I'm sure I ripped mine just a little bit...But I just made it a little bit looser." (Parent of an older infant)
Setting up and removal of LittleBeats	
Comfort with set up	• "I'm pretty comfortable getting it set up and turning it on. It seems pretty straightforward." (Parent of a school-aged child)
Child cooperation	• "It was mostly the initial putting it on. She didn't want to cooperate with letting us get it on...but after a little bit she forgot it was there because she didn't have any issues messing with it and then when it was time to take it off she was fine." (Parent of a preschool-aged child)
Location of device	• "I wish the device itself was a little more discreet. Well, he's a stomach sleeper so for naps I had to take it off but if it was a little more discreet or was not in front of the t-shirt but maybe on the arm it would be more convenient." (Parent of an older infant)
Participant burden	
Time commitment	• "You know once we broke it up a little bit we could [complete recordings]. I was more worried about you know were rarely all home just the four of us especially now that quarantine is over...We're just more on the go than we were a year ago." (Parent of an early school–aged child)
Convenience	• "It was really easy for me as a parent. I drive my other son like I said to [research lab in different city] a bunch...and so that is just a drag, a lot of back and forth. But for I would say from a parent's standpoint, this was very easy for me to do." (Parent of an early school–aged child)
Worry about recordings	• "My son's pretty active, so he sweats a lot over the course of the day. The little stickers would kind of migrate a little bit...So, I worry a little bit that the first recording like the second half of the day might not be as accurate as it was supposed to be." (Parent of a preschool-aged child) • "It would be nice if there were some kind of indicator of battery more visible. And it was also, you know, since I had to take it on and off then count the time, that was also kind of challenging...so some kind of indication of time would also be awesome but I don't know how complicated it would be to make it." (Parent of an older infant)
Worry about device	• "A lot of the activities that she wants to do involve painting or drinking water...those kinds of worrying me every time she picks them up. I was more concerned about the hardware." (Parent of a preschool-aged child)

Privacy

The privacy domain focuses on the types of personal information that are or will be collected about participants. In this study, privacy relates to participants' expectations about and understanding of the process of data collection, in general, and the audio recordings, specifically. Furthermore, this category encompasses the control that participants had over the data collected (Table 2).

Many parents commented that they were initially apprehensive about the home audio recordings but that their worries subsided when provided with more details during the initial informational call with the study coordinator. Other parents noted feeling more comfortable with the audio recordings over time as they participated in the study. Some parents discussed that although

XSL•FO
RenderX

they had no concerns, their spouse or partner did. Typically, only 1 parent (usually the mother) was present for the initial informational call with the study coordinator, and this parent then conveyed information to the other parent, which often sufficed to relieve privacy concerns.

By contrast, for some parents, positive views of research, such as having trust or placing value in research, negated concerns about privacy. Other participants described not being concerned with the audio recordings because they "had nothing to hide." From this view, the audio would capture a typical day in their life, and participants elaborated by describing that the recordings would include everyday family discussions as well as arguments, which participants conveyed as just part of ordinary family life. Others' lack of concern regarding the audio recordings stemmed from their ability to control when they were recording and, consequently, what was being recorded. They described the process of turning the device on and off as relatively easy and, therefore, reported turning the device off when they were discussing private matters. Some participants mentioned developing ground rules ahead of time to ensure that private information was not discussed when recordings were taking place and, if needed, would alert or remind other family members of the recordings.

The possibility of recording other individuals beyond immediate family members was considered. In working to respect others' privacy, the participants mentioned several challenges. Some participants expressed that they altered their typical day to avoid interacting with others so that they would not have to worry about unintentionally recording a nonconsenting individual. Other participants stated that although they had planned to record at convenient times when no nonconsenting individuals were around, unexpected situations arose. In addition, although parents had the ability to control when the device recorded, some parents acknowledged that remembering to turn off the device when others were around could be challenging.

Table 2. Themes, subthemes, and example excerpts related to privacy concerns about the LittleBeats audio recordings (study 1).

Themes and subthemes	Example excerpts
Initial apprehension about audio recordings	• "Cause that was my husband's big question like 'are they just going to sit and listen to our day?' So, he was a little worried about that but once it was explained [that machine learning algorithms would be used to analyze the audio data] he was more comfortable and on board." (Parent of an early school–aged child)
Unconcerned about audio recordings	
Former views or experiences of research	• "She [study coordinator] also told me that it is only used for research purpose and nothing else...I actually love to participate [in research studies]. It is only used for research purposes, so that's okay."(Parent of a preschool-aged child) • "I was in a study when I was pregnant and we did something similar...my understanding was that the recordings just gets run through the software so we really don't have anything any interesting happening here so I wasn't terribly concerned about that [the audio recordings]." (Parent of an older infant)
Just an ordinary family context	• "I explained everything to everybody [family members, including older children in home]. I do remember there was one particular situation where my 10-year-old was getting into trouble and afterwards he said, 'Well, they're gonna hear that!' And I said this is just a regular family, there's nothing to be embarrassed about or whatever." (Parent of a toddler)
Ability to control the recordings	• "My husband's a veteran, and he works at the V.A..., so we had to make sure we turned it off before he came home from work because a lot of times he talks about his day." (Parent of a toddler)
Respecting others' privacy	
Adjusting routines or activities to accommodate the study	• "I think the only thing is that we didn't go play with some friends across the street those days where we would've otherwise. Like it impeded a little bit of our typical routine, but it felt pretty unobtrusive." (Parent of a preschool-aged child)
Unexpected situations	• "When something was happening that I wasn't expecting, like when I would get a phone call or something like that, and I was just a little concerned about remembering to turn off the device." (Parent of an older infant)

Risks and Benefits

Evaluating the risks of possible harms in relation to the possible benefits resulting from the knowledge to be gained from the research is linked to the principle of beneficence. Study benefits should outweigh the possible harm to participants and the groups they represent. Risk assessment includes evaluating the type of harm, psychological, physical, reputational, or economic. In addition, researchers must consider the duration, severity, and intensity of the possible harm. Specific to the risks associated with the use of LittleBeats at home, parents expressed varying views along several dimensions, including safety, child comfort, and understanding of the research and its direct outcomes for participants (Table 3).

Many parents expressed that they thought the device was safe for their children to wear. These parents described not being concerned about safety because of the design of the device and the protective features built into it (eg, device was enclosed, tape-covered wires, fitted shirt, and pocket with secure snaps).

XSL•FO
RenderX

Some parents indicated that they initially had safety concerns (eg, the device being close to the skin and use of Bluetooth to transfer data) before learning more about the device and its setup (eg, the device itself is not in contact with the skin but is placed in a padded pocket, data are stored directly on the device, and Bluetooth is not used for data transfer). In some instances, parents detailed concerns about their children wearing the device in unsupervised contexts, such as during naptime, and they preemptively removed the device before naps.

Parents also commented on their children's level of comfort or discomfort. Several parents mentioned that they observed their child functioning normally, such as engaging in typical routines and activities. Parents also stated that their children did not express any discomfort and did not seem to notice that they were wearing the device after a while. Other parents noted their children's discomfort in putting on or removing the electrodes and medical tape used to secure the wires on the chest. Some parents worried about how comfortable it would be if the child were to hit the device on another object, such as the edge of a table.

Finally, parents' understanding of the research and its direct outcomes for their families may confer risks and benefits. Some parents revealed a limited understanding of how the data would be used (ie, the ultimate outcome of the research process) or wanted direct feedback on their children's development, which could pose unintended risks (eg, unfulfilled expectations of direct benefits). Other parents voiced the benefits attributed to participating in the research project itself. For instance, participation provided dedicated time spent together as a family, or completing the surveys was an opportunity to reflect on their children's activities and development. Several parents expressed their desire to contribute to the project because they recognized the importance of the research. Some parents indicated that they had enjoyed participating in previous studies, and others stated that this project's description seemed interesting and fun. Other parents of older children revealed that when they initially talked to their children about the study, their children seemed interested in participating, so they signed up. Some participants communicated that their children enjoyed participating in the project, with one parent acknowledging that their children felt special for a day while wearing the LittleBeats shirt.

Table 3. Themes, subthemes, and example excerpts related to the risks and benefits of participating in the LittleBeats study (study 1).

Themes and subthemes	Example excerpts
Safety	• "No [safety concerns] because all of the wires were covered by her shirt and taped down." (Parent of an older infant) • "Not really [any safety concerns]. I mean the wires were short enough that I wasn't worried about them." (Parent of an early school–aged child) • "I thought it will get like hot because I recorded for the 8 hours straight, I didn't stop it at all, I was worried maybe it's gonna be hot or something, but it wasn't hot at all. That was my main concern only." (Parent of a younger infant) • "And then I did have an initial concern...about the safety of having that device running on Bluetooth. I'm not sure how it communicates data and that being so close to skin." (Parent of an older infant)
Child's comfort or discomfort	• "I guess putting them [electrocardiogram electrodes] on wasn't the hard part. The hard part was taking them off, especially the was a little bit hard, and my son is also not very fond of changing clothes." (Parent of an older infant) • "I'd probably take it off especially because my little one is about 10 1/2 months and she's a tummy sleeper so that would be uncomfortable." (Parent of a preschool-aged child) • "I mean it seemed it was fine. My sons were playing outside you know riding their bikes and everything and they didn't...say anything was uncomfortable." (Parent of an early school–aged child)
Outcomes of participating in the research	
Limited understanding	• "I would love to know what kind of information. I know what kind of information they collected with the device and I'm just curious what they are going to use it for in the future." (Parent of an older infant)
Understanding gained	• "[Filling out] this survey, I found that I am pretty lucky that my son is more adaptable. The question, was for example, 'when you want him to go to bed, he just cried or tantrum' but he never does that." (Parent of a preschool-age child)
Parent's enjoyment or satisfaction	• "I just like participating in research and helping out the scholars. In my undergrad, I was doing some research and I know how important it is and how hard it can be so...I think it's good to help." (Parent of an older infant) • "I actually like to spend time with my son. He goes to school every day, so I like to do something with him like the zoom interview. And also I want to show him new technologies." (Parent of a preschool-aged child)
Child's enjoyment	• "I didn't mind the surveys or anything, and my son loved wearing the LittleBeats. He kept asking if he could put them on. So, I think it captured the kid's interest too." (Parent of an early school–aged child) • "We had fun doing it [the study], and I think [my son] enjoyed being special, wearing his special shirt for a day." (Parent of a toddler)

XSL•FO
RenderX

Study 2

Overview

Building on the key themes of access and usability, privacy, and risks and benefits identified in study 1, we administered a brief survey among a larger sample of parents participating in a different LittleBeats study with children aged 0 to 5 years. Although our main interest was to complement the qualitative findings of study 1 with a quantitative assessment of parents' perceptions using close-ended rating scales, parents were also able to provide open-ended comments. Therefore, we have also summarized the main themes reflected in these open-ended comments.

Methods

Participants

In study 2, a total of 110 parents (n=108, 98% mothers and n=2, 2% fathers) completed a user experience survey after completing 3 days of LittleBeats recordings at home. Recruitment procedures were similar to those described in study 1. Children (60/110, 54.5% female) were aged, on average, 23.4 months (SD 16.87 mo; range: 2-65 mo) and were identified by parents as Black (n=5, 4.7%), Asian (n=8, 7.5%), White non-Hispanic (n=67, 63.2%), Hispanic (n=15, 14.2%), or >1 race (n=11, 10.4%). Children were first born (n=50, 47%), second born (n=39, 38%), and third or later born (n=17, 15%). Parents were aged, on average, 34.85 (SD 5.01) years, and their highest level of education reported included some high school or high-school degree (4/106, 3.8%), some college or 2-year degree (9/106, 8.5%), a bachelor's degree (33/106, 31.1%), or an advanced degree (60/106, 56.6%). Parents identified as Black (7/106, 6.6%), Asian (13/106, 12.3%), White non-Hispanic (75/106, 70.8%), Hispanic (8/106, 7.5%), or >1 race (3/106, 2.8%). The mean family income was US $83,250 (SD US $26,470). Of the 110 parents, 4 (4%) were missing responses on the demographic survey but did complete the LittleBeats user experience survey described subsequently.

Ethical Considerations

This study was approved by the institutional review board at the UIUC (protocol #22631).

Procedure

Families were mailed a LittleBeats kit and participated in a Zoom visit, during which a study coordinator walked the parent through the LittleBeats setup and a visit procedure consisting of a baseline assessment of child stress physiology and parent-child interaction tasks (eg, play). At the end of the visit, parents received instructions about completing the day-long recordings and were asked to complete 3 day-long recordings over the course of 2 weeks. Parents also completed a series of web-based questionnaires about family demographics, child behavior, and family functioning. Parent questionnaires were administered either via Qualtrics or REDCap (Research Electronic Data Capture; Vanderbilt University [41,42]) hosted at the UIUC, with the support of the Interdisciplinary Health Sciences Institute and Research IT—Technology Services at the UIUC. Both web-based software platforms are designed to support secure data capture for research studies. Once parents returned the LittleBeats kit by mail, 1 parent in the household (who had been involved in setting up and carrying out the LittleBeats recordings) was asked to rate 5 items about their experience of using LittleBeats, including setting up LittleBeats, along with their perceptions of safety, privacy, and participant burden. Each item was rated on a 5-point scale ranging from 1 (*strongly agree*) to 5 (*strongly disagree*). Following each item, parents had the opportunity to add comments or elaborate on their rating. A final open-ended item also asked parents whether there was anything else they would like to share about their experience or anything they would tell someone who was considering joining a LittleBeats study.

Data Analytic Plan

Descriptive statistics, including the frequency distribution, for parental ratings on each of the LittleBeats user experience items were examined. For each close-ended item, we conducted a single-sample *t* test (2-tailed) to determine whether the mean rating significantly differed from the midpoint of the scale (ie, value of 3="neutral"). Finally, using the coding scheme developed in study 1, we assessed themes from parents' responses to the open-ended items.

Results

Parents' Ratings on User Experience Items

Percentage frequency distributions of parents' ratings on the user experience items are shown in Table 4. Single-sample *t* tests indicated a significant difference between the item average (lower ratings indicated greater agreement; higher rating indicated greater disagreement) and the midpoint of the rating scale (3="neutral") for 4 (80%) of the 5 items. Compared with a "neutral" response, parents were significantly more likely to *agree* that (1) the LittleBeats instructions were helpful and clear (mean 1.21, SD 0.41; t_{109}=−45.98; *P*<.001), (2) they felt comfortable setting up LittleBeats on their child (mean 1.42, SD 0.75; t_{109}=−22.22; *P*<.001), and (3) they felt their child was safe while wearing LittleBeats (mean 1.33, SD 0.51; t_{109}=−34.48; *P*<.001). Compared with a "neutral" response, parents were significantly more likely to *disagree* that they worried about being recorded by the LittleBeats device (mean 3.62, SD 1.06; t_{108}=6.14; *P*<.001). The final item tapped parents' perceptions of burden ("I felt that completing LittleBeats recordings for full 3 days was challenging"), and the item average (mean 2.98, SD 1.17) did not significantly differ from "neutral" (t_{109}=−.16; *P*=.87).

XSL•FO
RenderX

Table 4. Frequency distributions of parental rating of the LittleBeats user experience survey (study 2; n=110).

Survey item	Strongly agree, n (%)	Agree, n (%)	Neutral, n (%)	Disagree, n (%)	Strongly disagree, n (%)
The instructions to setup LittleBeats were helpful and clear.	87 (79.1)	23 (20.9)	0 (0)	0 (0)	0 (0)
I felt comfortable setting up LittleBeats on my child.	75 (68.2)	29 (26.4)	2 (1.8)	3 (2.7)	1 (0.9)
I felt my child was safe while wearing LittleBeats.	76 (69.1)	32 (29.1)	2 (1.8)	0 (0)	0 (0)
I was worried about being recorded when the LittleBeats device was on.	3 (2.7)	13 (11.8)	32 (29.1)	35 (31.8)	26 (23.6)
I felt that completing LittleBeats recordings for 3 full days was challenging.	11 (10)	32 (29.1)	26 (23.6)	30 (27.3)	11 (10)

Parents' Responses to Open-Ended Items

A review of parents' responses to the optional item to add further comments following each of the rating scales revealed themes that closely mirrored study 1 findings. Regarding the ease-of-use item, 29 (26.4%) of the 110 parents added comments. Most parents noted that having an instruction card included in the kit, as well as a QR code to easily link to the website for more detailed instructions, increased usability.

Regarding comfort in setting up the device, 23 (20.9%) of the 110 parents added comments. Parents noted that they felt comfortable and that the setting up of the device was easy. However, parents also noted that the process of setting up the device was difficult when their child moved around. Other parents mentioned the comfort level of their child (eg, noting that their child felt discomfort when removing the ECG electrodes).

Regarding safety, 25 (22.7%) of the 110 parents added comments. Parents noted few concerns because the device was concealed in a pocket and not easily accessible to the child. Parents who expressed a concern commented on the placement of the device on their child's chest.

Regarding concerns about being recorded, 32 (29.1%) of the 110 parents added comments. Some parents noted feeling self-conscious about their parenting or other family members' language choices. Typically, these comments were followed by comments about feeling relieved that the audio would be processed by a machine (vs a human coder). By contrast, many parents explained that they went about their day as usual, which typically contained some sort of sibling argument or other family disagreements.

Regarding participant burden, 68 (61.8%) of the 110 parents added comments. Unlike in study 1, where participants were asked to use the device for 2 days, study 2 participants were asked to use the device for 3 full days (or a total of about 24 hours) over the course of 2 weeks. Several parents commented on their families' busy schedules and difficulty finding 3 full days when only immediate family members were present.

Finally, a number of parents (46/110, 41.8%) responded to the final open-ended question asking whether they had any other comments they would like to share. Responses mirrored study 1 themes in several respects, including parents' and children's enjoyment in participating in the study (eg, "fun and easy" and "I would recommend to my friends"), children's ability to forget about the device and go about their usual day (eg, "did not interfere with our day"; "[Child] did not notice the device...he was able to nap with it on and so it was really pretty simple to participate!"; and "once the shirt was on, she forgot it was there and so did I!"), and suggestions for ways to minimize burden and improve the experience (eg, adding a display on the device that provides more information about battery charge, power status, and recording length).

Discussion

Summary

Digital health technologies have largely been developed with adults in mind. Interest in and attention to the use of wearable devices among infants and young children, however, has been growing, and data collection using wearable devices provides several advantages over traditional data collection methods, including continuous assessment, greater ecological validity, and the automated detection of behaviors using machine learning algorithms. Given these advantages, combined with rapid technological advances, it is likely that the use of wearables in child development research will burgeon in the coming years. Therefore, assessing how such devices and related data collection protocols are perceived and experienced by parents and their children is critical. User experience studies not only address ethical considerations but can also lead to important changes in research protocols that address parents' concerns and increase the benefits for future families who participate. Indeed, our mixed methods investigation across 2 studies yielded consistent findings that shed light on parents' experiences and perceptions of LittleBeats' usability and safety, the privacy of the audio recordings, and potential risks and benefits of participating in research of this kind. A large majority of parents indicated that device instructions were helpful and clear, the device was easy to use and safe, and remote visits were convenient. Parents' views about privacy, risks, and benefits were more varied, although, on average, parents reported feeling comfortable with the audio recordings. In summarizing the major themes identified within the major categories, we consider ways in which the findings can inform the future design and implementation of wearable platforms in child development research.

XSL•FO
RenderX

Key Findings

Results across all themes underscored the variability in parents' (mostly mothers') perspectives and experiences. With respect to access and usability, some parents expressed interest in having access to information that indicated the cumulative time recorded as well as the battery charge remaining. Such additions to the platform would eliminate parents' need to track the recording length and minimize parents' concerns about whether the device was sufficiently charged and recording. Some parents also noted difficulty with threading the ECG lead wires through the back of the shirt or were worried that their child would tug on the wires. These challenges can be remedied by changing the shirt design such that the ECG wires would be more fully integrated into the shirt fabric or design. Although parents indicated that day-long recordings (ie, >8 h/d) were feasible, some parents noted challenges. To alleviate the burden of day-long recordings, the time requirements can be adjusted to be more flexible. For instance, parents can be asked to complete recordings for fewer hours per day across multiple days (ie, 3 to 4 h/d across 4 to 5 d), although the optimal length and frequency of recordings needed to reliably capture the constructs of interest will vary as a function of the research questions being addressed. Importantly, such burdens were balanced by parents' comments regarding the convenience of remote visit procedures and the ease of using LittleBeats.

Privacy was a theme that also garnered a variety of responses. Some parents indicated few concerns about the privacy of the home audio recordings, whereas other parents worried that the recordings captured private conversations. In the latter case, some families used rules or reminders to control or limit when audio recordings were collected. It is also notable that parents within the same family sometimes expressed differing levels of comfort or concern with the audio recordings. When this pattern emerged, it was largely fathers who voiced concern about invasion of privacy, perhaps because they were not present for initial conversations with the study coordinator, who detailed how the data would be collected and used.

We consider 2 main ways to address parents' privacy concerns about the home audio recordings (also refer to the study by Cychosz et al [13]). First, providing specific and concrete examples of how the audio recordings are processed and analyzed, perhaps by illustrating a hypothetical example of the data collection, processing, and analysis steps, may help ease privacy concerns. Indeed, some parents noted that the use of machine learning algorithms to analyze the data alleviated their concerns about the audio recordings and privacy-related issues. Thus, describing the machine learning algorithms in a detailed yet accessible manner for nontechnical users and stating ways in which the data will not be used or analyzed (eg, no transcriptions of speech) may help reassure parents. Such information should be provided to all family members participating in the home recordings, including older siblings, and should be presented in various formats (eg, brief informational videos, hard copy pamphlets, interactive web page), along with multiple ways to contact study personnel for questions or comments. As part of this solution and building on some parents' perspectives that the recordings were just capturing "typical family life," researchers conducting day-long recordings may also explicitly highlight the family as an important context for development, coupled with appreciation for the fact that all families are different, and that, as researchers, we want to capture what life is like for each family and infant.

A second solution to alleviate parents' concerns about privacy could involve technological innovations, such as collecting audio recordings in which speech content is not intelligible (refer to the study by Levin et al [33]) or data processing (eg, machine learning algorithms) that occurs on the device or hub in the home so that the audio recordings are not stored or released to the researcher. However, these solutions require further technological advances in audio signal processing and raise issues regarding data-quality assurance. That is, without high-fidelity recordings, the validation and quality checks of machine learning algorithms become difficult. Furthermore, when parents were presented with several hypothetical scenarios for collecting child sensor data in the home environment, parent-reported willingness to participate did not significantly differ between study scenarios in which lower resolution audio data were collected (eg, recording 1-min snippets every 20 min and processing audio data automatically so that raw audio data are not stored) and study scenarios in which higher resolution data (eg, continuous audio recordings) were collected [33]. Taken together, although technological solutions aimed at increasing privacy protection seem to be a reasonable avenue to pursue, future studies on users' experiences of child wearables, particularly home audio or video recordings, should systematically assess parents' concerns, needs, and desires when it comes to balancing the privacy of day-long home recordings with the benefits of participation.

Third-party or bystander privacy is also a complex issue [37,38]. In this study, there were two categories of potential third parties: (1) nonparental caregivers or relatives at home who were part of the child's regular routine and (2) individuals who were not part of the home environment (eg, delivery persons and neighbors). In the first case, nonparental caregivers can be included in the recording if they provided consent. In the second case, the parent would need to turn off the device while the individual is present or change their routine to avoid third parties, which may have consequences for ecological validity. Concerns about third-party recordings can also be resolved by the same types of technical solutions outlined earlier.

The principle of beneficence yielded a variety of responses regarding the risks and benefits of the study procedures. First and foremost, safety was a key theme, and across both samples, parents predominantly expressed views that LittleBeats was safe. When concerns about safety were mentioned, parents often presented hypothetical concerns (eg, the device being close to the skin, the device radiating heat, and the child accidentally falling on the device; the last scenario is mentioned as a potential risk in the parental consent form), which were usually alleviated once the parent learned more about the study. Some parents also mentioned concerns about the child wearing the device during unsupervised times, such as naps, and removed the device during these times. Because infants and young children are much more likely to take ≥1 naps over the course of the day, this subtheme differed across age groups, with parents of children in younger age groups being more likely to mention

XSL•FO
RenderX

device use with respect to nap times. Another set of risks is related to the child's discomfort, particularly around the application and removal of the ECG electrodes. This potential risk is also mentioned in the parental consent form, and we aimed to ameliorate this risk using latex-free electrodes designed specifically for pediatric populations.

Potential or perceived risks were balanced by parents' perceived benefits, including increased understanding of their child's development through the completion of the parent surveys, parents' satisfaction in contributing to the scientific process, children's enjoyment of the study procedures (eg, play session with parents), and wearing the novel LittleBeats shirt and device. We note that we did not ask directly about perceived benefits in study 2 close-ended items, although parents in this study did indicate the benefits of participation in the final open-ended question asking whether they had any other comments they would like to share. These responses often paralleled the positive sentiments that study 1 parents expressed. Nevertheless, items that assess the perceived benefits of study participation will be important to include in future studies.

With respect to increasing direct benefits to participants, we gave families personalized books summarizing information that we have collected about their children (eg, height and weight at different ages) in prior studies. Such summaries have been well received and appreciated. Similar types of summaries can be made from data extracted from day-long recordings (eg, frequency and duration of infant babbling or crying). Providing this type of study feedback to parents may also promote effective participant recruitment and retention, particularly among studies that involve high-fidelity data, such as audio recordings. As noted by Levin et al [33], individuals are likely to evaluate intrusiveness and data privacy, on the one hand, and direct benefits to themselves and their children (such as receiving useful, personalized information or feedback from the data collected), on the other hand, when making decisions about whether to participate in such research.

Study Limitations and Future Directions

We note several limitations of our user experience studies. First, we did not ask our older child participants about their experiences directly, although parents reported on a variety of child experiences, including compliance with putting on the device, excitement in wearing the shirt, feeling special while wearing the shirt, and comfort or discomfort. The device hardware was relatively compact and lightweight, and parents reported that children tended to forget about it once it was on. Nonetheless, these reflections clearly highlight the need to directly assess not only parents' perspectives but also children's perspectives. Thus, parental reports of their child's experiences should be augmented by direct observations of infants and younger children while wearing the device as well as interviews with older children. Second, we tracked parents' reported experiences based on the child's developmental stage. Similar themes were found across developmental periods, although specific examples of how themes manifested often differed by the child's age. However, because the subsamples of children in different age groups were relatively small, future research with larger subsamples is needed to more thoroughly investigate developmental considerations related to user experiences in the context of research using child wearables. However, an age-specific consideration that did clearly emerge relates to daytime sleep. Third, in both samples, parents reported high levels of educational attainment. Future research on parents' perspectives of using child wearable devices in the home setting should include families with diverse demographic characteristics. Including samples characterized by sociodemographic factors in user experience studies is especially critical for child wearables developed for the purposes of mobile health interventions.

Conclusions

Wearable sensors designed for and validated with infants and young children present researchers and clinicians with tremendous opportunities to assess developmental processes and outcomes in more ecologically valid and potentially less burdensome ways than laboratory assessments. Furthermore, LittleBeats' multiple modalities provide especially rich data to assess an array of constructs central to child development researchers and clinicians, including parent-child vocal turn-taking, regulation of stress, sleep-wake cycles, physical activity, and developmental disorders. At the same time, although we have validated LittleBeats sensors and machine learning algorithms to accurately capture some of these key constructs [17-19,43], the degree to which LittleBeats and similar child wearables deliver benefits (eg, high ecological validity and low burden) will largely depend on acceptance by the end users (eg, parents and children), making user experience studies critical to this research space. In short, if the technology is not acceptable to the end user, it is less likely to be adopted and used as intended. The user experience assessment presented in this paper goes hand in hand with technical validations of the device, and both are critical for successful implementation. The current results suggest that parents predominantly view LittleBeats as easy to set up and use at home, although views regarding privacy and burden were more varied. On the basis of parents' thoughtful and specific feedback, several concrete changes can be implemented to improve the LittleBeats platform and, ultimately, parents' and children's experiences.

Acknowledgments

The authors are grateful to the families who participated in this research and to Jenny Baldwin for her assistance with participant recruitment and data collection. This work was supported by funding from the National Institute on Drug Abuse (R34DA050256), National Institute of Mental Health (R21MH112578), and USDA National Institute of Food and Agriculture (ILLU-793-368).

XSL•FO
RenderX

Conflicts of Interest

None declared.

Multimedia Appendix 1
Parent user experience interview.
[DOCX File , 16 KB - humanfactors_v11i1e49316_app1.docx]

References

1. Peake JM, Kerr G, Sullivan JP. A critical review of consumer wearables, mobile applications, and equipment for providing biofeedback, monitoring stress, and sleep in physically active populations. Front Physiol 2018 Jun 28;9:743 [FREE Full text] [doi: 10.3389/fphys.2018.00743] [Medline: 30002629]
2. Sawka MN, Friedl KE. Emerging wearable physiological monitoring technologies and decision aids for health and performance. J Appl Physiol (1985) 2018 Feb 01;124(2):430-431 [FREE Full text] [doi: 10.1152/japplphysiol.00964.2017] [Medline: 29097633]
3. Muhammad G, Alshehri F, Karray F, Saddik AE, Alsulaiman M, Falk TH. A comprehensive survey on multimodal medical signals fusion for smart healthcare systems. Inf Fusion 2021 Dec;76:355-375 [FREE Full text] [doi: 10.1016/j.inffus.2021.06.007]
4. Mohr DC, Zhang M, Schueller SM. Personal sensing: understanding mental health using ubiquitous sensors and machine learning. Annu Rev Clin Psychol 2017 May 08;13(1):23-47 [FREE Full text] [doi: 10.1146/annurev-clinpsy-032816-044949] [Medline: 28375728]
5. Jaworski DJ, Park A, Park EJ. Internet of things for sleep monitoring. IEEE Instrum Meas Mag 2021 Apr;24(2):30-36 [FREE Full text] [doi: 10.1109/mim.2021.9400950]
6. Fuller D, Colwell E, Low J, Orychock K, Tobin MA, Simango B, et al. Reliability and validity of commercially available wearable devices for measuring steps, energy expenditure, and heart rate: systematic review. JMIR Mhealth Uhealth 2020 Sep 08;8(9):e18694 [FREE Full text] [doi: 10.2196/18694] [Medline: 32897239]
7. Harari GM, Lane ND, Wang R, Crosier BS, Campbell AT, Gosling SD. Using smartphones to collect behavioral data in psychological science: opportunities, practical considerations, and challenges. Perspect Psychol Sci 2016 Nov 28;11(6):838-854 [FREE Full text] [doi: 10.1177/1745691616650285] [Medline: 27899727]
8. Seifert A, Hofer M, Allemand M. Mobile data collection: smart, but not (yet) smart enough. Front Neurosci 2018 Dec 18;12:971 [FREE Full text] [doi: 10.3389/fnins.2018.00971] [Medline: 30618590]
9. Stewart E, Milton A, Yee HF, Song MJ, Roberts A, Davenport T, et al. eHealth tools that assess and track health and well-being in children and young people: systematic review. J Med Internet Res 2022 May 12;24(5):e26015 [FREE Full text] [doi: 10.2196/26015] [Medline: 35550285]
10. Memon SF, Memon M, Bhatti S. Wearable technology for infant health monitoring: a survey. IET Circuits Devices Syst 2020 Jan 29;14(2):115-129 [FREE Full text] [doi: 10.1049/iet-cds.2018.5447]
11. Airaksinen M, Räsänen O, Ilén E, Häyrinen T, Kivi A, Marchi V, et al. Automatic posture and movement tracking of infants with wearable movement sensors. Sci Rep 2020 Jan 13;10(1):169 [FREE Full text] [doi: 10.1038/s41598-019-56862-5] [Medline: 31932616]
12. Wang Y, Williams R, Dilley L, Houston DM. A meta-analysis of the predictability of LENA™ automated measures for child language development. Dev Rev 2020 Sep;57:100921 [FREE Full text] [doi: 10.1016/j.dr.2020.100921] [Medline: 32632339]
13. Cychosz M, Romeo R, Soderstrom M, Scaff C, Ganek H, Cristia A, et al. Longform recordings of everyday life: ethics for best practices. Behav Res Methods 2020 Oct 26;52(5):1951-1969 [FREE Full text] [doi: 10.3758/s13428-020-01365-9] [Medline: 32103465]
14. Nuske HJ, Goodwin MS, Kushleyeva Y, Forsyth D, Pennington JW, Masino AJ, et al. Evaluating commercially available wireless cardiovascular monitors for measuring and transmitting real-time physiological responses in children with autism. Autism Res 2022 Jan;15(1):117-130 [FREE Full text] [doi: 10.1002/aur.2633] [Medline: 34741438]
15. Lee R, Skinner A, Bornstein MH, Radford AN, Campbell A, Graham K, et al. Through babies' eyes: practical and theoretical considerations of using wearable technology to measure parent-infant behaviour from the mothers' and infants' view points. Infant Behav Dev 2017 May;47:62-71 [FREE Full text] [doi: 10.1016/j.infbeh.2017.02.006] [Medline: 28347907]
16. de Barbaro K. Automated sensing of daily activity: a new lens into development. Dev Psychobiol 2019 Apr 18;61(3):444-464 [FREE Full text] [doi: 10.1002/dev.21831] [Medline: 30883745]
17. Li J, Hasegawa-Johnson M, McElwain NL. Towards robust family-infant audio analysis based on unsupervised pretraining of Wav2vec 2.0 on large-scale unlabeled family audio. arXiv Preprint posted online on May 21, 2023 [FREE Full text] [doi: 10.21437/interspeech.2023-460]
18. Chang KC, Hasegawa-Johnson M, McElwain NL, Islam B. Classification of infant sleep/wake states: cross-attention among large scale pretrained transformer networks using audio, ECG, and IMU data. arXiv Preprint posted online on June 27, 2023 [FREE Full text] [doi: 10.1109/apsipaasc58517.2023.10317201]

XSL•FO
RenderX

19. Islam B, McElwain NL, Li J, Davila M, Hu Y, Hu K, et al. Preliminary Technical Validation of LittleBeats™: A Multimodal Sensing Platform to Capture Cardiac Physiology, Motion, and Vocalizations. Sensors 2024 Jan 30;24(3):901. [doi: 10.3390/s24030901]
20. Nebeker C, Torous J, Bartlett Ellis RJ. Building the case for actionable ethics in digital health research supported by artificial intelligence. BMC Med 2019 Jul 17;17(1):137 [FREE Full text] [doi: 10.1186/s12916-019-1377-7] [Medline: 31311535]
21. Nebeker C, Bartlett Ellis RJ, Torous J. Development of a decision-making checklist tool to support technology selection in digital health research. Transl Behav Med 2020 Oct 08;10(4):1004-1015 [FREE Full text] [doi: 10.1093/tbm/ibz074] [Medline: 31120511]
22. The Belmont report: ethical principles and guidelines for the protection of human subjects of research. National Commission for the Protection of Human Subjects of Biomedical and Behavioral Research. 1979. URL: https://www.hhs.gov/ohrp/regulations-and-policy/belmont-report/read-the-belmont-report/index.html [accessed 2023-12-14]
23. Dittrich D, Bailey M, Maughan D. The Menlo report: ethical principles guiding information and communication technology research. IEEE Secur Priv 2012 Mar;10(2):71-75 [FREE Full text] [doi: 10.2139/ssrn.2445102]
24. Nebeker C, Zlatar ZZ. Learning from older adults to promote independent physical activity using mobile health (mHealth). Front Public Health 2021 Aug 12;9:703910 [FREE Full text] [doi: 10.3389/fpubh.2021.703910] [Medline: 34476229]
25. Behnke M, Saganowski S, Kunc D, Kazienko P. Ethical considerations and checklist for affective research with wearables. IEEE Trans Affect Comput 2022:1-14 [FREE Full text] [doi: 10.1109/taffc.2022.3222524]
26. Nebeker C, Gholami M, Kareem D, Kim E. Applying a digital health checklist and readability tools to improve informed consent for digital health research. Front Digit Health 2021 Jul 15;3:690901 [FREE Full text] [doi: 10.3389/fdgth.2021.690901] [Medline: 34713167]
27. Thompson RA. Vulnerability in research: a developmental perspective on research risk. Child Dev 1990 Feb;61(1):1-16 [FREE Full text] [doi: 10.1111/j.1467-8624.1990.tb02756.x]
28. Tavakoulnia A, Lakes K, Guzman KD, Hayes GR, Cibrian FL, Schuck SE. Designing a wearable technology application for enhancing executive functioning skills in children with ADHD. In: Proceedings of the 2019 ACM International Joint Conference on Pervasive and Ubiquitous Computing and Proceedings of the 2019 ACM International Symposium on Wearable Computers. 2019 Presented at: UbiComp/ISWC '19; September 9-13, 2019; London, UK p. 222-225 URL: https://dl.acm.org/doi/10.1145/3341162.3343819 [doi: 10.1145/3341162.3343819]
29. Díaz-Quesada G, Bahamonde-Pérez C, Giménez-Egido JM, Torres-Luque G. Use of wearable devices to study physical activity in early childhood education. Sustainability 2021 Dec 18;13(24):13998 [FREE Full text] [doi: 10.3390/su132413998]
30. Bagot KS, Matthews SA, Mason M, Squeglia LM, Fowler J, Gray K, et al. Current, future and potential use of mobile and wearable technologies and social media data in the ABCD study to increase understanding of contributors to child health. Dev Cogn Neurosci 2018 Aug;32:121-129 [FREE Full text] [doi: 10.1016/j.dcn.2018.03.008] [Medline: 29636283]
31. Nakib M, Hasan U, Negulescu II. Wearable technology for baby monitoring: a review. J Text Eng Fash Technol 2020 Jul 14;6(4):112-120 [FREE Full text] [doi: 10.15406/jteft.2020.06.00239]
32. Geršak V, Vitulić HS, Prosen S, Starc G, Humar I, Geršak G. Use of wearable devices to study activity of children in classroom; case study — learning geometry using movement. Comput Commun 2020 Jan;150:581-588 [FREE Full text] [doi: 10.1016/j.comcom.2019.12.019]
33. Levin HI, Egger D, Andres L, Johnson M, Bearman SK, de Barbaro K. Sensing everyday activity: parent perceptions and feasibility. Infant Behav Dev 2021 Feb;62:101511 [FREE Full text] [doi: 10.1016/j.infbeh.2020.101511] [Medline: 33465730]
34. Nebeker C, Ellis RB, Torous J. Digital health checklist for researchers (DHC_R). ReCODE Health. 2019. URL: https://recode.health/tools/ [accessed 2024-01-04]
35. Health tools are licensed under a creative commons attribution-non-commercial 4.0 international license 2018-2024. ReCODE. URL: https://recode.health/tools/ [accessed 2024-01-08]
36. Zhang D, Lim J, Zhou L, Dahl AA. Breaking the data value-privacy paradox in mobile mental health systems through user-centered privacy protection: a web-based survey study. JMIR Ment Health 2021 Dec 24;8(12):e31633 [FREE Full text] [doi: 10.2196/31633] [Medline: 34951604]
37. Kelly P, Marshall SJ, Badland H, Kerr J, Oliver M, Doherty AR, et al. An ethical framework for automated, wearable cameras in health behavior research. Am J Prev Med 2013 Mar;44(3):314-319 [FREE Full text] [doi: 10.1016/j.amepre.2012.11.006] [Medline: 23415131]
38. Nebeker C, Lagare T, Takemoto M, Lewars B, Crist K, Bloss CS, et al. Engaging research participants to inform the ethical conduct of mobile imaging, pervasive sensing, and location tracking research. Transl Behav Med 2016 Dec;6(4):577-586 [FREE Full text] [doi: 10.1007/s13142-016-0426-4] [Medline: 27688250]
39. Rampin R, Rampin V. Taguette: open-source qualitative data analysis. J Open Source Softw 2021 Dec 10;6(68):3522 [FREE Full text] [doi: 10.21105/joss.03522]
40. Braun V, Clarke V. Using thematic analysis in psychology. Qual Res Psychol 2006 Jan;3(2):77-101 [FREE Full text] [doi: 10.1191/1478088706qp063oa]
41. Harris PA, Taylor R, Minor BL, Elliott V, Fernandez M, O'Neal L, REDCap Consortium. The REDCap consortium: building an international community of software platform partners. J Biomed Inform 2019 Jul;95:103208 [FREE Full text] [doi: 10.1016/j.jbi.2019.103208] [Medline: 31078660]

XSL•FO
RenderX

42. Harris PA, Taylor R, Thielke R, Payne J, Gonzalez N, Conde JG. Research electronic data capture (REDCap)--a metadata-driven methodology and workflow process for providing translational research informatics support. J Biomed Inform 2009 Apr;42(2):377-381 [FREE Full text] [doi: 10.1016/j.jbi.2008.08.010] [Medline: 18929686]
43. Li J, Hasegawa-Johnson M, McElwain NL. Analysis of acoustic and voice quality features for the classification of infant and mother vocalizations. Speech Commun 2021 Oct;133:41-61 [FREE Full text] [doi: 10.1016/j.specom.2021.07.010] [Medline: 36062214]

Abbreviations

ECG: electrocardiogram
REDCap: Research Electronic Data Capture
UIUC: University of Illinois Urbana-Champaign

Edited by C Jacob; submitted 24.05.23; peer-reviewed by A Angelucci, J Wang; comments to author 04.08.23; revised version received 26.09.23; accepted 29.09.23; published 08.02.24.

Please cite as:
McElwain NL, Fisher MC, Nebeker C, Bodway JM, Islam B, Hasegawa-Johnson M
Evaluating Users' Experiences of a Child Multimodal Wearable Device: Mixed Methods Approach
JMIR Hum Factors 2024;11:e49316
URL: https://humanfactors.jmir.org/2024/1/e49316
doi:10.2196/49316
PMID:38329785

©Nancy L McElwain, Meghan C Fisher, Camille Nebeker, Jordan M Bodway, Bashima Islam, Mark Hasegawa-Johnson. Originally published in JMIR Human Factors (https://humanfactors.jmir.org), 08.02.2024. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on https://humanfactors.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Usability and User Experience of an mHealth App for Therapy Support of Patients With Breast Cancer: Mixed Methods Study Using Eye Tracking

Carolin Anders[1*], MSc; Preetha Moorthy[2*], MSc; Laura Svensson[1], MSc; Julia Müller[1], MSc; Oliver Heinze[1], Dr sc hum; Petra Knaup[1], Dr sc hum; Markus Wallwiener[3], Dr med; Thomas M Deutsch[3], Dr med; Thao-Vy Le[3]; Lina Weinert[1,4], MSc

[1]Institute of Medical Informatics, Heidelberg University Hospital, Heidelberg, Germany

[2]Department of Biomedical Informatics at the Center for Preventive Medicine and Digital Health (CPD), Medical Faculty Mannheim, Heidelberg University, Mannheim, Germany

[3]Department of Obstetrics and Gynecology, Heidelberg University Hospital, Heidelberg, Germany

[4]Section for Oral Health, Heidelberg Institute of Global Health, Heidelberg University Hospital, Heidelberg, Germany

[*]these authors contributed equally

Corresponding Author:
Carolin Anders, MSc
Institute of Medical Informatics
Heidelberg University Hospital
Im Neuenheimer Feld 130.3
Heidelberg, 69120
Germany
Phone: 49 622156 ext 34651
Email: Carolin.Anders@med.uni-heidelberg.de

Abstract

Background: Early identification of quality of life (QoL) loss and side effects is a key challenge in breast cancer therapy. Digital tools can be helpful components of therapeutic support. *Enable*, a smartphone app, was used in a multicenter, prospective randomized controlled trial in 3 breast cancer centers. The app simultaneously serves as a therapy companion (eg, by displaying appointments), a tool for documenting QoL (eg, by enabling data collection for QoL questionnaires), and documentation of patient-reported side effects. The need for digital tools is continually rising. However, evidence of the effects of long-term use of mobile health (mHealth) apps in aftercare for patients with breast cancer is limited. Therefore, evaluating the usability and understanding the user experience of this mHealth app could potentially contribute valuable insights in this field.

Objective: A usability study was conducted to explore how patients with breast cancer receiving neoadjuvant, adjuvant, or palliative outpatient treatment rated their engagement with the app , the user experience, and the benefits of using the app.

Methods: A mixed methods approach was chosen to combine subjective and objective measures, including an eye-tracking procedure, a standardized usability questionnaire (mHealth App Usability Questionnaire), and semistructured interviews. Participants were surveyed twice during the study period. Interviews were transcribed verbatim and analyzed using thematic analysis. Analysis of the eye-tracking data was carried out using the tracker-integrated software. Descriptive analysis was conducted for the quantitative data.

Results: The mHealth App Usability Questionnaire results (n=105) indicated good overall usability for 2 different time points (4 wk: mean 89.15, SD 9.65; 20 wk: mean 85.57, SD 12.88). The qualitative analysis of the eye-tracking recordings (n=10) and interviews (n=16) showed that users found the *Enable* app easy to use. The design of the app, information about therapies and side effects, and usefulness of the app as a therapy companion were rated positively. Additionally, participants contributed requests for additional app features and suggestions for improving the content and usability of the app. Relevant themes included optimization of the appointment feature, updating the app's content regularly, and self-administration. In contrast to the app's current passive method of operation, participants expressed a desire for more active engagement through messaging, alarms, or emails.

Conclusions: The results of this study demonstrate the good usability of the *Enable* app as well as the potential for further development. We concluded from patients' feedback and requests that mHealth apps could benefit from giving patients a more

XSL•FO
RenderX

active role (eg, being able to actively document side effects as they occur). Additionally, regular updates of app content could further contribute to encouraging continued use of mHealth apps. Our findings may also assist other researchers in tailoring their mHealth apps to the actual needs of patients undergoing breast cancer therapy.

(JMIR Hum Factors 2024;11:e50926) doi:10.2196/50926

KEYWORDS

mobile health; mHealth; usability; breast cancer; eye tracking; user interface; mixed methods; mobile phone

Introduction

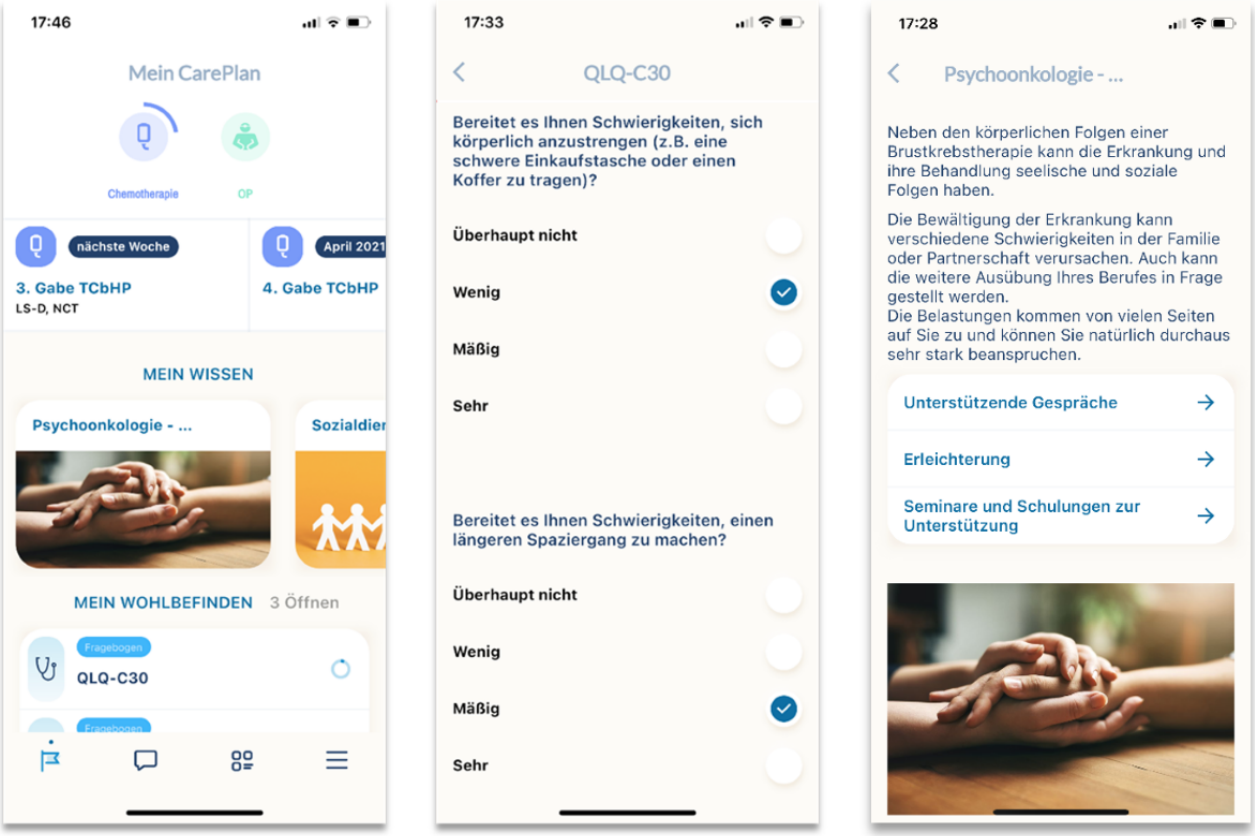
Background

Breast cancer is the most common type of cancer detected in women in the Western world. One in 8 women will develop breast cancer during her lifetime. In Germany, there are 69,000 new cases per year [1]. The diagnosis is a drastic event in the lives of those affected. Although the mortality rate has decreased in recent years, processing and dealing with the new life situation is a great challenge for patients and their social environment [2]. At the onset of therapy, patients can have a strong desire for education and information. Therefore, providing patients with reliable sources of information and support services is a major and important task for the treatment team. Digitalization in medicine offers great potential for supporting the exchange of information and communication between patients and health care providers [3-5]. These benefits can be realized through the use of mobile health (mHealth) apps, which can encompass several helpful functions for patients, such as the provision of educational materials, appointment or medication reminders, and diaries. For the cohort of patients with breast cancer, many of these mHealth apps are already available or are in development [6]. This cohort also shows a high readiness for using health technology, indicating that mHealth apps are an appropriate means of support in the early phase of breast cancer treatment.

A recent study by Chen et al [7] also found that remote monitoring of symptoms between clinical visits could not only improve patient-provider communication but also prepare patients for subsequent chemotherapy cycles and support symptom management. Within the joint *Center for Innovative Care* project, a network of 5 university hospitals in southwest Germany, a new mHealth app for patients with breast cancer was developed. This therapy support tool, called the *Enable* app, aims to combine known benefits of mHealth tools with an innovative reactive assessment of patient-reported outcomes (PROs). It was conceptualized as an iOS or Android mobile app for smartphones and developed by members of the research team with the support of software developers. It includes educational content, information about the side effects of therapies and medications, and information about other support services such as psycho-oncology or nutritional counseling in the form of static text and images. A progress bar illustrates the patient's individual therapy status in terms of clinical treatment over time (ie, cycles of treatment). In addition to its role as a therapy companion, the app serves as a measurement tool to systematically record patient satisfaction, health-related quality of life (QoL), and patient-reported adverse events. It monitors the neoadjuvant, adjuvant, and follow-up situations in patients with indications for surgery, chemotherapy, radiation, or systemic therapy with primary or metastatic breast cancer. Figure 1 shows exemplary screenshots of the *Enable* app's start page, the questionnaire display, and information about treatments. As studies have shown that physicians generally underestimate a large proportion of relevant side effects, patients are empowered to report PRO data and side effects directly through the app. In cases of significant treatment-related deterioration, the care team is alerted, and recommendations are sent to the patient. This more relevant treatment information, in turn, helps improve therapy monitoring, treatment quality, and patient satisfaction [8,9].

XSL•FO
RenderX

Figure 1. Exemplary content of the Enable app. Left: On the start page of the app, patients can see their current therapy status, a selection of articles from the "My Knowledge" collection, and upcoming questionnaires; Middle: View of the QLQ-C30 questionnaire to assess the quality of life of cancer patients; Right: Sample article on psycho-oncology with the aim of patient education.



The clinical outcomes of the use of the *Enable* app were studied in the *ENABLE* randomized controlled trial (RCT). Other research questions addressed in the *ENABLE* RCT related to improving patients' adherence to therapy, recognizing and treating critical side effects in a timely manner, and measuring the health-related QoL of different therapy strategies. All study participants underwent QoL assessments at 6 time points during and after adjuvant or neoadjuvant chemotherapy. In the intervention group, an additional short weekly EuroQol Visual Analogue Scale questionnaire was administered. In case of deteriorating results, further screening for side effects was triggered, alerting study staff and enabling immediate contact with the patient to provide support in all phases of breast cancer therapy (reactive PRO assessment). The control group received only the app without the reactive PRO assessment.

The body of scientific literature shows that good usability is an important factor for the success of an mHealth app. More specifically, usability can influence patients' acceptance and adoption of mHealth [10,11]. Usability is defined by Nielsen [12] as a "quality attribute that assesses how easy interfaces are to use." According to the International Organization for Standardization (ISO) 9241-1, usability is the "extent to which a system, product, or service can be used by specified users to achieve specified goals with effectiveness, efficiency, and satisfaction in a specified context of use" [13]. Although usability focuses exclusively on the process of using an app or device, user experience involves the users' subjective feelings that result from the use or anticipated use of a system or a product. For the evaluation of mHealth, both concepts are relevant to obtain a comprehensive view of influencing factors [14].

Good usability can help ensure that the app can be used intuitively by patients and health care providers, which in turn improves compliance and increases the effectiveness of the app. A review by Zapata et al [10] demonstrated the importance of adapting mHealth apps to patients' needs. Relevant usability themes of similar apps were, for example, streamlining of the navigation paths, a clearer information architecture, or the desire for personalization [15,16]. Recent research has also shown that usability assessment is an essential step in the mHealth app development process [17,18]. It is important to ensure that the app is easy to use for the target group and provides the desired benefits [12]. However, a systematic review by Jongerius et al [6] showed that only 1 of 29 mHealth apps for breast cancer care that were studied in their work underwent and published a usability assessment. To address the aforementioned requirements and achieve sustainable and effective use of the *Enable* app, the investigation of usability and user experience is indispensable. Therefore, the study presented in this paper intended to gain an understanding of how patients use the app. The aim was to investigate how patients evaluate their engagement with the app, the user experience, and the benefits of using the app. These findings will serve as a basis for further optimization and adaptation of the app to the patients' needs.

A mixed methods approach provides the opportunity to collect, triangulate, and analyze qualitative and quantitative data,

XSL•FO
RenderX

allowing for the possibility of interpreting the findings from one research approach (ie, qualitative and quantitative) to explain the data generated from the other research approaches. Furthermore, it allows for the use of a qualitative approach to illustrate quantitative findings or the integration of various research approaches to provide a thorough and comprehensive picture of the study [19,20]. Previous studies [15,21,22] have indicated that interviews and usability questionnaires are prevailing methods used for assessing the usability of mHealth apps. However, there are limited studies regarding the real-time capture of users' visual interactions and the subsequent retrospective analysis of user engagement with mHealth apps through techniques such as eye tracking. Eye tracking, a sensor technology, is used to ascertain an individual's presence and record their real-time eye movements. This approach is also used to assess the usability of technologies by showcasing decision-making processes through the analysis of eye movement patterns [23,24].

Objectives

Developing new mHealth apps can be time-consuming and requires several iterations of testing and evaluation. The *ENABLE* project aims to evaluate both the usability and clinical outcomes of the *Enable* app within the same RCT, which could be a promising approach to speed up development, testing, and planning for further implementation. This paper presents a usability study nested within the *ENABLE* RCT and following a mixed methods approach incorporating the eye-tracking method. The objective of this usability study was to explore how patients with breast cancer receiving neoadjuvant, adjuvant, or palliative outpatient treatment rated their engagement with the app, the user experience, and the benefits of using the app.

Methods

Study Design

This study was designed following a mixed methods approach combining real-world user experience and standardized observations in a laboratory setting. The study took place at the Department of Obstetrics and Gynecology, Heidelberg University Hospital, Germany.

Procedure

Study Population and Recruitment

The study participants were recruited from the intervention and control groups of the *ENABLE* RCT patient cohort (German Clinical Trials Register—DRKS ID: DRKS00025611). The *ENABLE* RCT had the following inclusion criteria: diagnosis of invasive or metastatic breast cancer and planning of neoadjuvant, adjuvant, or palliative therapy in an outpatient treatment setting (indications for surgery or chemo-, radio-, or systemic therapy); minimum age of 18 years; German language skills; and possession of a smartphone with internet access. Owing to technical requirements for eye tracking, patients wearing bifocals were excluded from participation. At study enrollment, patients were asked about their interest in participating in the usability study. All interested patients at the Department of Obstetrics and Gynecology, Heidelberg University Hospital, Germany, received written and verbal information regarding the content and aim of the study and the respective data protection regulations. On the informed consent form, patients could indicate whether they were interested in participating in the usability aspect of the *ENABLE* RCT. Patients who consented to participate in the nested usability study were contacted individually to schedule appointments for participation following a convenience sampling strategy. No reimbursement was provided. The target sample size was 100 questionnaires, 15 qualitative interviews, and 10 eye-tracking studies. Patient recruitment took place from March 2021 to September 2023.

Instruments

The German translation of the mHealth App Usability Questionnaire (MAUQ) [25] was chosen to quantitatively assess the usability of the *Enable* app [26]. The MAUQ enables the usability assessment of mHealth apps from the user's perspective. The MAUQ stand-alone version was formulated to evaluate 3 constructs of usability—ease of use, interface and satisfaction, and usefulness—as well as the overall usability score for the app through descriptive statistics. Each of the items of the MAUQ is rated on a Likert scale ranging from 1 (*strongly agree*) to 7 (*strongly disagree*), with the overall score ranging from 0 to 100. In addition, the questionnaires were complemented with a set of questions developed by the authors. Newly added questions concerned the use of other mHealth apps, smartphone ownership, sociodemographic information, and a free-text field to be able to describe the study sample more precisely. The target sample size was 100.

In addition to the questionnaire, open-ended, semistructured, and guide-based interviews with patients were conducted to explore their perspectives on the usability of the *Enable* app. The interviews were conducted by 2 female researchers (CA and LW) with a professional background in health services research and implementation science. Both researchers have profound experience with qualitative interviewing. The interview guide (Multimedia Appendix 1) was developed by a team of health services researchers (LW and JM) based on an extensive literature review and recommendations from the app developers. Afterward, the interview guide was pretested. This study is reported according to the COREQ (Consolidated Criteria for Reporting Qualitative Research) guidelines (Multimedia Appendix 2 [27]).

Furthermore, to objectively assess how patients interact with the app and identify potential usability issues, an eye-tracking study was conducted. The eye-tracking study was conducted by a usability expert (PM) and a team of health services researchers (CA and LW). A total of 5 tasks were formulated for the eye-tracking study (Multimedia Appendix 3): app log-in, filling in a questionnaire, searching and reading an article, and logging out from the app. To determine the comprehensibility of the tasks, the duration of the study, and the workings of the *Enable* app, 2 pilot tests were conducted. Following the pilot test outcome, the eye-tracking studies were carried out for 60 minutes with each participant, including the eye tracker setup and the retrospective interview.

The chosen mixed methods approach is designed to systematically collect, cross-validate, and analyze both

XSL•FO
RenderX

qualitative data (derived from semistructured interviews and eye tracking) and quantitative data (obtained through the MAUQ). The inclusion of the eye-tracking method in the usability study enriches the capacity to integrate subjective and objective metrics. The qualitative aspect of the eye-tracking analysis enhances the understanding of the user's app perception within the context of individual interactions and app usability. Simultaneously, semistructured interviews enable an assessment of the practicality of integrating the *Enable* app into daily routines. In contrast, the quantitative data derived from the questionnaire provide precise metrics related to usability measurements.

Hence, the mixed methods approach investigates the *why* and *how* aspects through qualitative inquiry, supplementing conventional quantitative and visual data analyses. The fusion of direct observations of user interactions with the app, poststudy retrospective interviews, semistructured interviews, and the usability questionnaire collectively supports the contextualization and comprehensive interpretation of the gathered data.

Data Collection and Analysis

Quantitative Measures

The MAUQ and sociodemographic questionnaire were mailed twice to all patients after inclusion in the RCT. Data collection lasted from May 2021 to October 2022. Study data were collected and managed using REDCap (Research Electronic Data Capture; Vanderbilt University) tools [28] hosted at Heidelberg University Hospital. REDCap is a secure, web-based software platform designed to support data capture for research studies. After completion, all data were exported from REDCap to the R statistical software (version 4.0.4; R Foundation for Statistical Computing). All data were checked for completeness and analyzed by study team members. A descriptive analysis of the questionnaires was performed using R. Means and absolute and relative frequencies were calculated.

Qualitative Measures

Interviews were conducted after participants had used the app for 8 weeks. The interviews took place partly face-to-face at the clinic and by telephone in consideration of current guidelines for preventing infections with SARS-CoV-2 (ie, participants and researchers wore appropriate masks and distance was kept at all times). Nonparticipants were not present during the interviews. No relationship with participants was established before taking part in the study. No repeated interviews were conducted. No field notes were taken. All interviews were audiotaped, pseudonymized, and transcribed verbatim. Transcripts were not returned to participants for verification. Data were transcribed, managed, and analyzed using MAXQDA Standard 2020 (version 20.4.1; VERBI GmbH). After 16 interviews, data saturation was discussed among the researchers. As no new themes emerged in later interviews, the researchers agreed that data saturation had been reached and no additional interviews were necessary. After completion of data collection, thematic analysis of the data was conducted independently by 2 researchers (CA and LW) [29]. First, the researchers reviewed the transcripts independently and identified themes from the literature and the interview guide and inductively from the data. Second, discrepancies were discussed in iterative cycles until a consensus on themes and the final coding scheme was reached. All themes were organized into main themes and subthemes. Each theme was clearly defined by a quote from the interview transcripts (Multimedia Appendix 4). Quantitative and qualitative data were analyzed separately.

For the eye-tracking data collection process, an assigned room where the Tobii Pro Nano (Tobii AB) was installed at the hospital was used; the Tobii Pro Nano is an eye-tracking device specifically designed for small screens, including smartphones. This hardware features a sampling rate of 60 Hz, measures 17 × 1.8 × 1.3 cm, and includes a USB type-A connector. The Tobii Pro Nano was securely affixed to the mobile phone stand, and the *Enable* app was installed on a smartphone. To facilitate data capture, both the smartphone and the eye tracker were connected to a laptop running the Tobii Pro Lab software (version 1.194) via USB cables. For the purposes of this study, both an Android device (Samsung Galaxy 10, Android version 11) and an iOS device (iPhone 11, iOS version 14.6) were available to users. The choice of smartphone was contingent upon the user's preferred operating system. The eye tracker recorded the participants' interactions with the *Enable* app, such as task completion time, participants' navigation, gaze plots, and heat maps [30-32]. A heat map was used when fixation duration data were collected [30,31], and a gaze plot was used when location of eye movement data were collected [33,34]. For this study, after the completion of tasks, the study moderators composed post hoc questions pertaining to the interactions, participants' experiences, and usability issues observed during the procedure. The post hoc questions were discussed with the participants in a short debrief. The debriefing sessions were held to gather direct feedback from participants after interacting with the *Enable* app, allowing for a deeper understanding of the participants' behavior and interaction with the app. Through these debriefing sessions, participants could provide context and commentary on their behavior and interaction [35]. Engaging users using post hoc questions, such as using images or live content from recorded sessions, allowed for a better understanding of the real-life context with minimal disruption as it facilitated the recall of situational information prompted by data, sound, or visual imagery.

The data analysis was based on the recordings of the study sessions concurrent with the eye movements of participants. The retrospective analysis involved transcribing participants' feedback from the audio recordings obtained during the debriefing sessions. Data analysis also included the completion of predefined tasks by the participants, task completion time, and completion status of the tasks. The analysis focused on task performance analysis and the problem analysis of eye-tracking metrics and participants' feedback.

Ethical Considerations

The study was conducted in accordance with the Declaration of Helsinki and approved by the Ethics Committee of Heidelberg University Hospital (S-685/2020). All participants provided written informed consent for taking part, audio recording of the interviews, and video recordings during the eye-tracking

XSL•FO
RenderX

procedures. Confidentiality and anonymity were ensured throughout the study. The data was protected against unauthorized access. No incentives or compensation was provided to participants for study participation.

Results

Overview

The MAUQ was sent to 165 patients recruited from the *ENABLE* RCT. The response rate was 63.6% (105/165) for the MAUQ at week 4 and 56.4% (93/165) for the MAUQ at week 20. A total of 105 questionnaires for the MAUQ at week 4 (including sociodemographic data) and 93 questionnaires for the MAUQ at week 20 were analyzed. In total, 16 patients were recruited for the interviews, and 10 were recruited for the eye-tracking procedure. The mean duration of the interviews was 25 (SD 7.34) minutes.

Sociodemographic Characteristics

The sociodemographic data of the participants in the *ENABLE* usability study are shown in Table 1, and additional characteristics of the participants regarding smartphone and app use are shown in Table 2. The mean age of all participants (n=105) was 51.3 (SD 10.9) years.

Table 1. Sociodemographic characteristics of the participants.

Characteristic	Interview participants (n=16), n (%)	Eye-tracking study participants (n=10), n (%)	Questionnaire participants (n=105), n (%)
Gender			
Woman	16 (100)	10 (100)	105 (100)
Age group (y)			
<30	2 (12.5)	2 (20)	2 (1.9)
30-40	2 (12.5)	1 (10)	16 (15.2)
41-50	6 (37.5)	3 (30)	32 (30.5)
51-60	4 (25)	2 (20)	33 (31.4)
61-70	1 (6.3)	0 (0)	16 (15.2)
71-80	1 (6.3)	0 (0)	6 (5.7)
Education			
Academic degree	9 (56.3)	6 (60)	37 (35.2)
High school education	0 (0)	0 (0)	13 (12.4)
Lower or intermediate secondary school	5 (31.3)	4 (40)	54 (51.4)
Prefer not to say	2 (12.5)	0 (0)	1 (1)
Employment			
Employed	11 (68.8)	9 (90)	66 (62.9)
Unemployed	0 (0)	0 (0)	17 (16.2)
Studying or vocational training	1 (6.3)	1 (10)	1 (1)
Retired	2 (12.5)	0 (0)	18 (17.1)
Prefer not to say	2 (12.5)	0 (0)	3 (2.9)

XSL•FO
RenderX

Table 2. Additional participant characteristics on smartphone and app use.

Characteristic	Interview participants (n=16), n (%)	Eye-tracking study participants (n=10), n (%)	Questionnaire participants (n=105), n (%)
Use of smartphone (y)			
≤10	9 (69.2)[a]	5 (62.5)[b]	56 (53.3)
>10	4 (30.8)[a]	3 (37.5)[b]	44 (41.9)
Prefer not to say	0 (0)	0 (0)	5 (4.8)
Use of other mHealth[c] apps			
Yes	4 (28.6)[d]	5 (50)[e]	33 (31.4)
No	10 (71.4)[d]	5 (50)[e]	71 (67.6)
Prefer not to say	0 (0)	0 (0)	1 (1)
Frequency of app use			
Daily or several days a week	5 (45.5)[f]	3 (37.5)[b]	0 (0)
Once a week	5 (45.5)[f]	4 (50)[b]	0 (0)
Once a month or less	1 (9.1)[f]	1 (12.5)[b]	0 (0)

[a]n=13.
[b]n=8.
[c]mHealth: mobile health.
[d]n=14.
[e]n=10.
[f]n=11.

Quantitative Measures

The MAUQ [25] was used to collect quantitative data on the usability of the *Enable* app. The data were collected at weeks 4 and 20 starting from the baseline of the study. Quantitative data gathered from the MAUQ were analyzed using descriptive statistics. Only complete questionnaires for which the MAUQ score could be calculated were evaluated. Hence, 32.4% (34/105) of incomplete questionnaires collected at week 4 and 29% (27/93) of incomplete questionnaires collected at week 20 were excluded from the analysis. According to Zhou et al [25], the usability of an app is calculated based on the average of the responses to all statements. The higher the overall average, the higher the usability of the app. In this study, the overall usability scores for weeks 4 and 20 were 89.15 (SD 9.65) and 85.57 (SD 12.88), respectively. The mean for each of the subscales from week 4 to week 20 was also calculated and is presented in Table 3. The results show that the usefulness score declined over time from week 4 (80.89) to week 20 (77.33). In addition, the *interface and satisfaction* score also decreased but not as much as that of the *usefulness* subscale. The *ease of use* score, in contrast, remained constant at both weeks 4 and 20.

Table 3. Quantitative analysis of the mHealth App Usability Questionnaire and subscales.

Time point	Overall, mean (SD)	Ease of use, mean (SD)	Interface and satisfaction, mean (SD)	Usefulness, mean (SD)
Wk 4 (n=71)	89.15 (9.65)	92.41 (11)	91.6 (10.15)	80.89 (15.67)
Wk 20 (n=66)	85.57 (12.88)	92.27 (11.91)	88.23 (13.6)	77.33 (16.63)

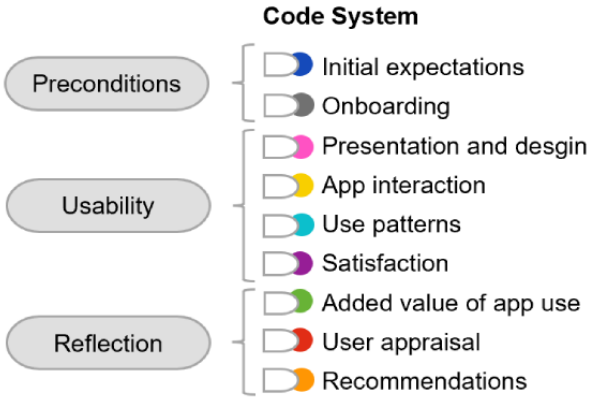
Qualitative Measures

Interviews

In total, 527 text passages were coded during the interviews. A total of 9 themes and 60 subthemes were identified, each of which could still be categorized under the superordinate themes of preconditions for app use, usability, and reflection. These themes are summarized in Figure 2.

XSL•FO
RenderX

Figure 2. Overview of identified themes.



Preconditions

Initial Expectations

As an opening question in the interview, patients were asked about their initial thoughts when they first heard about the *Enable* app. The most frequently mentioned expectations were related to the quality of information. Patients expected the information in the app to be updated regularly, understandable, and in line with the latest research. Another expectation was that the app would provide contemporary therapy support and be perceived as modern, including replacing printed brochures. Patients expected the app to provide guidance over the course of therapy, contact options, and easy access to relevant information. Approximately half of the participants had neutral expectations for the app:

> *Yes, I already thought that it [the app] would support me through everyday life and therapy, that I can also use it to organize myself a bit.* [Interview 9; transcript position 2]

Onboarding

The aspect of the onboarding process was not part of the interview guide. However, individual participants reported that they felt well supported by the study staff at the beginning of their app use. Even if they were initially overwhelmed by the app or experienced technical difficulties, participants expressed that they received the necessary support and were able to handle the app:

> *Oh dear, now I have to dig into yet another app. I don't know if I can handle it. But the more I got a grip on it, the better it worked.* [Interview 3; transcript position 2]

Usability

Presentation and Design

Patients were asked to describe their impressions of specified design aspects. Overall, patients were content with the color scheme and perceived it as pleasant without being boring or flashy. For some patients, this cheerful esthetic contributed to a sense of joy when using the app and encouraged them to use it more often:

> *[The design is] very friendly. Very beautifully visualized. I always enjoy opening the app. It is also well designed, you always have the feeling that it is not draining in any way, it is more playful with all these images and visualizations. I find it very very clear.* [Interview 11; transcript position 34]

Regarding the layout of the written information, patients appreciated how the most relevant parts were highlighted through the positioning of boxes. The font size, design, and structuring of the information were seen as adequate. The selection of accompanying images was described as empathetic and not too explicit. The app included a personalized visual representation of the therapy progress. This display was also rated as clear and useful. Patients explained that the presence of this display motivated them:

> *I found this progress bar, which shows me how long I will be in therapy for, especially beautiful. It...motivated me, showing me that there is always a path forward and that the therapy will soon be over.* [Interview 13; transcript position 43]

App Interaction

Regarding usability, 4 important aspects emerged while interacting with the app. Neither the positioning nor the design of the app icons were perceived as entirely intuitive. However, patients grew acclimated to the icons, and thus, this did not further impede usability:

> *Yes, the icons that were down in this bar. In the beginning, I didn't know the meaning of each icon. But when I took a closer look once, I knew it for the next time.* [Interview 7; transcript position 39]

Log-in and log-out procedures were described as easy and quick and did not pose any problems for the patients in this study. Most patients had no issues working with the app's structure. They could easily navigate within the app and were able to find what they were looking for:

> *I found my way around the app really quickly. I haven't tried all the features yet, I haven't clicked on everything because I don't need it all. But I have always been able to find the things that I wanted very quickly, and everything is right there when you click on it.* [Interview 11; transcript position 44]

XSL•FO
RenderX

Overall, patients liked using the app as it enabled them to access information *on the go*. Patients described that having their smartphones with them at all times allowed them to read information given the absence of other electronic devices such as laptops or tablets. However, a few patients mentioned the additional benefits of having a web-based version of the *Enable* app.

Use Patterns

This code encompasses descriptions of how and when patients used the app. Most patients experienced changes in the frequency of app use. In the beginning, they used the app often, and some patients used it multiple times per day:

> *In the beginning, shortly after my diagnosis, I had a lot of questions—for my physicians, how things work and so on. During this time, it (the app) really helped me a lot.* [Interview 10; transcript position 24]

Over time, use declined. This development was mostly due to lower demand for support and information as patients became used to therapy proceedings. Patients also used the app less as they felt that they had already read everything.

After this initial phase, patients reported using the app whenever they needed to look up appointments, had free time (eg, during waiting times before physician's appointments), had or experienced new side effects from their treatment, were prescribed new medications, or were prompted by push notifications:

> *I always used it shortly before my [chemotherapy] appointments. Or when I had questions regarding diet and exercise. And sometimes there were questionnaires I had to fill in. And yes, as soon as the app said "there is news," I opened it...And to look up times for my appointments.* [Interview 10; transcript position 12]

Satisfaction

Patients praised the general aspects of the app and liked the idea of having a digital tool accompanying them throughout their therapy; for example, the app provides a good overview of relevant topics, especially at the beginning of the disease. Except for 1 interviewee, all participants (15/16, 94%) would recommend the app to others:

> *...because it really provides a great overview...because so many aspects are addressed. Not only the type of therapy, but also just different things about cancer. Especially at the beginning these keywords—Yes, these terms in the boxes from tiredness to fatigue and polyneuropathy and different things.* [Interview 7; transcript position 47]

Reflection

Added Value of App Use

When asked about the concrete benefits of the app in everyday life, several aspects were mentioned. The most important aspect for the participants was the information on therapies and side effects, which was perceived as helpful, especially in the initial phase of therapy. The quality of the information was praised as the app's information was considered understandable and its origin was considered reliable:

> *You feel informed, you feel—that gives you a form of security, because you say to yourself: Well, if I have the information from here [the app], then it was completely clear to me: I don't have to look it up again. That's true for me because these are reliable information providers who wrote this.* [Interview 12; transcript position 81]

The comprehensibility and language level were also perceived as adequate. Statements on the amount of information were heterogeneous according to individual information needs. However, the amount of information was predominantly perceived as sufficient in the context of the app. Furthermore, the appointment display, contact information, and progress bar were found to be helpful and clear. With regard to the contact information provided in the app, the fact that it was easy to find was rated positively.

Some patients reported that the questionnaires in the app gave them a positive feeling as they reflected on their condition and (in the intervention group) it was experienced positively that the questionnaires were read by the study staff and that staff could react proactively to them if necessary. Overall, patients perceived the app as a good therapy companion that guided and supported them through the various phases of the disease and therapies.

User Appraisal

Users' opinions on the existing functions and features of the app were added to this category. Most patients complained about the appointment display as the date and time on the app did not always correspond to the actual clinic appointments (eg, in the case of last-minute postponements):

> *It's a shame that the—I don't know how the appointments displayed in the app, how often those are matched. I've had frequent differences there. Especially when appointments had to be postponed, the chronology was no longer correct for me.* [Interview 9; transcript position 2]

Regarding the quantity of information, some patients wished for more in-depth information or links to other information platforms. It was remarked that the amount of information available varied depending on the topic. Regarding the quality of the content, patients noted that the listed side effects or drugs were grouped differently. For instance, the patients were unable to locate paclitaxel as it belonged to the taxane drug class. In total, 12% (2/16) of the patients in particular perceived errors in spelling, punctuation, and grammar as distracting. The presentation of the contact information on the app was described as difficult to find, especially in emergencies. The additional pop-up notifications of the app updates were rated negatively as it was not apparent to the user what exactly was new in the app. Furthermore, respondents ascertained that the menu navigation was not intuitive enough and, therefore, needed to be improved.

XSL•FO
RenderX

Recommendations

Statements about features of the app that are not yet offered were classified as recommendations or wishes. Most wishes were mentioned in relation to the appointment display. Patients would like to have additional information about appointments, such as directions, a reminder function, the ability to export appointments from the app to private calendars (eg, Google or Outlook calendars), or the ability to make appointments directly from the app. The desire for self-administration (ie, areas such as appointments, questionnaires, or therapy progress that can be actively managed by the patient) was also frequently voiced. In addition, some patients wished to view the questionnaires that had already been completed to be able to monitor their condition over the course of therapy:

> *With the exception of filling in the questionnaires, you can't work with the app yourself. Therefore, if you could manage things in the app by yourself, then of course I would think that would be great.* [Interview 9; transcript position 2]

Patients also wanted the content of the app to be updated, expanded, and adapted to new scientific findings. In this context, there was a desire for more explanatory videos to be included in the app. Patients also suggested that the app should offer more information about current and upcoming clinical trials for patients with breast cancer. To see what content in the app has already been read, patients suggested a read status, where content that has already been read is highlighted. Emergency contacts should also be highlighted in the app to make them easier to find, for example, by displaying them on the home page:

> *Especially the emergency numbers, I don't know how to get something like that into the app, but that might be an idea, because I've been looking a lot for the right contact person. Maybe that would also be something that you could highlight a little bit or display as a button.* [Interview 4; transcript position 32]

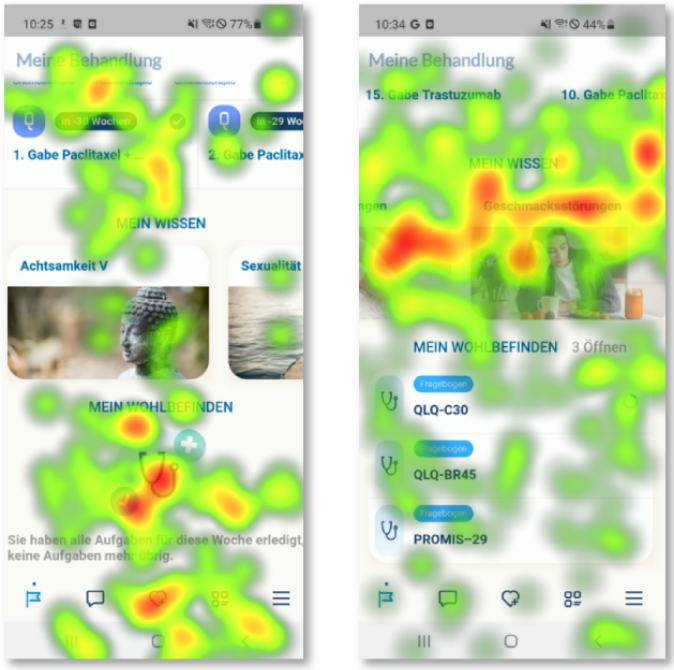
To be able to find certain topics more quickly, the need for a search function was mentioned several times. Furthermore, to improve the readability of the content, patients would like to be able to adjust the font size. It was also suggested that the app could be used on other devices, such as tablets.

Eye Tracking

Overview

The analysis of the data collected from the eye-tracking recordings as well as the retrospective interviews showed that the participants found the app easy to use. We observed that most participants completed the given tasks, although the time taken to complete a few tasks proved to be challenging. On the observations and retrospective interviews during the eye-tracking study, we discovered 3 noticeable patterns related to the design and layout of the app, content and navigation through the app, and additional features the participants would like to have in the app. Figure 3 shows exemplary heat maps from the eye-tracking analysis. The data collected during the task performance, such as the task completion rate and task completion times, are provided in Multimedia Appendix 5.

Figure 3. Heat maps from the eye-tracking analysis.



Design and Layout of the App

Many of the participants had problems understanding and interpreting the icon at the bottom of the screen. The eye-tracking data showed fixations at the bottom of the screen while the patients clicked each of the icons displayed to view the content of the page. Patients expressed a preference for finding the most important information, such as appointment

XSL•FO
RenderX

dates and the progress of a questionnaire, at the top of the screen. This finding indicates that patients expect important information to be located at the top of the app's layout. Furthermore, the patients actively mentioned that the retrievability and visibility of the questionnaire were low. Although the questionnaires were available on the home screen of the app, patients believed that the questionnaires were available on the menu. In contrast, patients found the overall layout of the app to be acceptable.

Content and Navigation of the App

Regarding the content of the app, patients showed more interest in the titles of the articles (eg, topics such as symptoms or side effects) than in the images displayed. When asked during the retrospective interviews, patients mentioned that they did not pay attention to the images as they provided no information on what the article was about. Patients preferred to read the title of the article as it gave them information about its content, as shown by the red areas of the heat maps in Figure 3. Moreover, many participants explored the app to find the right information or icon to perform the tasks. However, this correlates with how frequently patients used the app. During the interviews, some patients said that they used the app frequently, for example, every day, to read articles on side effects or symptoms and fill out questionnaires regularly, whereas some patients used the app frequently at the start phase of the *ENABLE* RCT and later minimized the use of the app except to fill out questionnaires. The data showed that patients also had issues navigating through the app, especially related to the task of finding a specific article. Analysis of the recorded data of the participants' navigation and gaze plots from the Tobii Eye Tracker showed that patients looked for a search function. Most patients clicked the menu icon; however, they did not proceed further to find the article nested under the *Symptoms* category on the menu. In addition, some patients searched for the article on the start page along with the other articles already displayed.

"Would Like to Have" (Wishes)

Participants identified a need for additional features in the *Enable* app as a consequence of the challenges they encountered during the eye-tracking study tasks. These suggested features were considered as *nice-to-have* options and were based on the specific problems faced by the participants during the study. The first was the availability of an option to mark an article as a favorite and be able to view the favorite article on the start page. Second, patients desired to have more articles or information about the symptoms and side effects of breast cancer and its treatments. Third, the icon currently representing contact information for health care providers (*My Care Team*) was misleading. Patients preferred to have another icon that indicates contact or communication as this would enable them to contact the study nurses more quickly. Finally, a search option was suggested by all participants.

Discussion

Principal Findings

The aim of this study was to investigate how patients with breast cancer rated their engagement with the *Enable* app, the user experience, and the benefits of using the app. In particular, the design, layout, navigation, content, and requests for new features were identified as important outcomes of interest for evaluating the app and further improving it to meet user needs. The interviews provided valuable suggestions for optimizing the app and the implementation process. The design and color scheme were rated very positively overall. In terms of use patterns, it was noticeable that the frequency of app use decreased over the therapy period.

Patients found the app easy to navigate. However, there was some criticism that the menu icons were not intuitive enough, especially at the onset of use. Perceived benefits were discussed extensively in the interviews. Patients found the information on therapies and side effects very useful. The appointment display and progress bar were also found to be helpful and motivating. At the same time, the appointment display was most often criticized, and it was the feature for which there were the most recommendations for change (eg, to be able to manage appointments autonomously in the app or set reminders). In terms of content, it was mentioned that there was a lot of information on some topics and not enough on others. Patients also wanted more content updates within the app (eg, on current topics such as the COVID-19 pandemic) and a search function to access specific content.

A study by Ansaar et al [36] showed that nearly 78% of all usability evaluation studies in their systematic review used a questionnaire-based method. However, using mixed methods approaches in usability evaluation studies provides benefits such as the possibility to balance the advantages and disadvantages of the different methods. Moreover, by applying the mixed methods approach, both subjective and objective aspects can be combined to assess usability [36]. In many aspects, such as the *navigation*, *recommendations*, and *perceived benefits* codes, the results of the different survey methods support each other. However, the interviews and eye-tracking study sometimes provided different findings. For example, the importance of images within the app was positively highlighted in the interviews. In contrast, the eye-tracking study and retrospective interviews revealed that images played a subordinate role for patients, with titles being more important for finding relevant content in the app. Although participants reported in the interviews that they were able to navigate easily within the app and find the content they were looking for, we observed in the eye-tracking study that there were difficulties with finding specific content. Furthermore, the interview inquiries primarily centered on the practicality of incorporating the *Enable* app as a follow-up intervention in daily life. Meanwhile, the use of eye-tracking technology allowed for direct, real-time observation of user behavior while engaging with the app through task performance. Despite patients reporting the ability to regularly use the app without difficulty, the eye-tracking study's direct observation unveiled valuable insights into their actual use patterns within their everyday routines. In this context, disparities between the results obtained from the 2 methods emerged, possibly stemming from users' lack of awareness regarding any issues until they were prompted with specific inquiries.

XSL•FO
RenderX

Comparison With Prior Work

Our results on the MAUQ indicate good usability. The results for the total scale showed that usability decreased from weeks 4 to 20. A decrease in usability over time has also been observed in previous studies [37-39]. Possible explanations for this decline in our study can be found in the interviews, indicating that the extent of app use also decreased over the course of therapy. Patients found the app to be particularly advantageous at the start of their therapy because of their great need for information. However, as they gained more knowledge about the disease and its treatment, their demand for information decreased. In addition, patients reported that the app lost its appeal once all the available articles had been read, often leading to a desire for new content to be added. Patients also expressed a need for additional features or improvements as they continued to use the app. As a result, the decrease in the app's usability score could be attributed to patients perceiving it to be less useful after an extended period of use owing to the lack of content updates and unmet desires.

Looking more closely at the subscales of the MAUQ, *usefulness* had the lowest score compared with *ease of use* and *user interface and satisfaction*. These items assess whether the app is helpful and useful for patients' health and well-being. This relationship is also apparent when looking at the *usage patterns* category from the interview analysis. It appears that patients are less likely to use the app because of the lack of new content. This is consistent with the findings of other studies on mHealth apps for patients with breast cancer [16,40,41]. As an implication for similar apps for other chronic conditions, it seems important to update the app content on a regular basis to provide patients with an incentive to continue using the app as well as strengthening patients' satisfaction and information needs. Consistent with the findings from the interviews and eye-tracking study, only the *ease of use* subscale remained almost stable over the duration of app use.

In the context of other usability studies on mHealth apps, the importance of paying more attention to the user group of older adults is emphasized. The different age ranges of patients and the different levels of technical affinity for older patients are mentioned as possible factors causing usability problems. Some studies emphasize that these factors are often overlooked and need to be considered when developing mHealth apps [42,43]. In our study, these aspects were less evident. With an average age of 51 years, our study participants do not represent a predominantly older population but are close to the German population average for women, which is 46 years [44]. In contrast, the study participants were also far below the average age of 64 years for patients with breast cancer. Therefore, further research on app development and usability with a focus on older participants should be conducted to more adequately represent the typical population of patients with breast cancer.

Considering the preferred device for using the *Enable* app, most participants were content with using the app on their smartphones. However, there were isolated requests to be able to increase the font size of the content and use the app on a larger-scale device, such as a tablet or PC. This issue was also mentioned by participants in a usability study by Jessen et al [45], in which an mHealth app for self-management of chronic diseases was evaluated.

Although the onboarding process was not part of the interview guide, some patients actively recalled how they were introduced to the app as well as how they perceived the technical onboarding process. The patients did not experience issues with these steps and reported being content with the process, mostly because of the strong support of the study team. Previous research has pointed out that complex registration and log-in procedures can be perceived as especially cumbersome by patients and can lead to stopping app use [46-48]. Our study identified the strong interpersonal connection with and continued support from the study team as a positive influence on the perceived ease of onboarding. This support took place in the context of a research study and is not viable in a real-world implementation. However, the issue of technical support arose exclusively during the qualitative interviews. We did not collect any quantitative data on this topic. Thus, further streamlining of the onboarding process while being mindful of health care workers' limited time resources should be an area for future research.

Strengths and Limitations

The chosen mixed methods approach can positively support the further development of the app. The expansion of the classic social science method spectrum to include technical methods such as eye tracking made it possible to combine the subjective patient perceptions reported in interviews and questionnaires during everyday use with objective measurements under laboratory conditions.

However, the integration of qualitative results and the objective measurement from the eye-tracking procedure introduced discrepancies. As noted previously, interviewees appreciated the use of images in the app, whereas eye-tracking results showed that more time was spent on the article titles than on the images. Another example is that the interviews and the questionnaire produced good ratings of usability, but the eye-tracking study showed that patients found it difficult to find defined content. Although difficult to analyze, these discrepancies are common in mixed methods studies [19]. In our study, these discrepancies could be explained by methodological differences. For example, reading a title naturally takes longer than glancing at an image, leading to a long fixation time. Therefore, this result does not allow for the conclusion that titles are more important than images. Here, the qualitative interviews were helpful in interpreting this finding. Regarding the second example—overall good usability scores in comparison with eye-tracking times—several interpretations appear plausible. First, it is possible that social desirability led patients to rate the usability more favorably in both the interviews and the questionnaire. Consequently, the objective measure via eye tracking revealed that usability was worse than in subjective measures. Second, the setting of the eye-tracking procedure (eg, unusual or uncomfortable sitting position, being observed by ≥2 researchers, or using a different device) could have led to changed patterns in (app use) behavior. Although we acknowledge these discrepancies, we conclude that the mixed methods approach and its results deepened the understanding

XSL•FO
RenderX

of the studied topic and produced valuable insights, with discrepancies leading to vigorous and fruitful discussions among the researchers.

However, the generalizability of the study results is limited by several factors. To ensure that patients with lower digital health literacy could participate in the quantitative data collection without constraints, we decided to use printed surveys sent by mail. Patients returned them at their discretion. Hence, it cannot be verified whether the surveys were filled out at the correct time. In addition, some values were missing from the returned surveys, and manual data entry could have led to documentation errors. Incomplete or inconclusive questionnaires had to be completely excluded from the analysis as it was not possible to calculate the score. Although all necessary steps were taken to ensure high-quality and reliable data (eg, data entry was always checked by another researcher), using a web-based survey instead of a printed survey could have made data collection easier, faster, and more reliable. These trade-offs have to be balanced in future research projects.

This study population contained an above-average proportion of academics, especially among the subgroups of interviewees and eye-tracking study participants. This should be taken into account when interpreting these results. A systematic review by Niazkhani et al [49] showed that patients with lower educational attainment and limited health literacy were less likely to intend to use an electronic patient health record and were more likely to use it ineffectively. Moreover, previous experience with computers or health technology has been associated with increased acceptance, and acceptance increases with higher education [7]. Although these results refer to electronic health records, they indicate that this aspect should be further investigated in future studies. Given the median age at breast cancer diagnosis of 64 years and the relatively younger median age of this study cohort, conclusions from this study must be interpreted with caution as they may not represent the views and digital literacy of older women with breast cancer [50].

The *Enable* app was developed specifically for patients with breast cancer. Consequently, our study sample included only female patients with breast cancer. Some of our results and recommendations may have limited generalizability to other patient populations. Nevertheless, we think that aspects such as the relevance of content updates, the accuracy of displayed appointments, or the intuitiveness of the app navigation might also be relevant beyond the target group. This should be verified in further research.

As part of the *ENABLE* RCT, reasons for dropping out were documented where available. These reasons were examined to see whether there were any indications of usability problems. A small proportion of the included study participants in the RCT dropped out because of physical exertion or feelings of being overwhelmed by the app. In this respect, further research is needed to understand how patients in later stages of the disease or with greater disease burden perceive the usability and benefits of the intervention. Furthermore, mHealth apps should be designed to be usable and helpful for these patient groups as well, especially in the context of patients living with cancer. As the mean age of participants in this study was relatively low, it can be assumed that there is a risk of selection bias. It is possible that younger patients decided to participate in the study and use the app because of a higher affinity for smartphones [11].

In addition, using the eye-tracking device led to further limitations. Potential participants in the eye-tracking study had to undergo an additional screening process to exclude patients wearing bifocal glasses. Although patients were recruited for the study, this criterion did not allow us to cast a wider net for the participant recruitment process. Furthermore, we also had the challenge of asking patients to sit still so that the eye-tracking data could be captured without breaks. However, this request is generally against the natural way in which users sit and interact with mobile devices. Another point to note is that the execution of the tasks on the app by the patients was deviated as the tasks were presented on paper and this retracted some of the gaze points of the patients. This is, in general, a common problem when tasks are not integrated into mobile apps during development for testing purposes.

Conclusions

The results of this usability study demonstrate good usability of the studied app and potential for purposeful development. The design and color scheme were rated very positively overall. However, there was some criticism that the menu icons were not intuitive enough, especially at the onset of use. Noticeably, the frequency of app use decreased over the therapy period. Perceived benefits of the app were information on therapies and side effects. The appointment display and progress bar were also found to be helpful and motivating. Still, participants offered recommendations for changing the appointment display (eg, to be able to manage appointments autonomously in the app or set reminders). In terms of content, it was mentioned that there was a lot of information on some topics and not enough on others. Patients also wanted more content updates within the app (eg, on current topics such as the COVID-19 pandemic) and a search function to access specific content. The interviews and eye-tracking study revealed valuable suggestions for improvement as well as requests for additional app features. An important point is that the app currently provides information to the patient mainly passively. The patients' wishes indicate that the app needs to be further developed so that they can actively enter information into the app and work with it. The overlap between decreasing usability and decreasing usefulness also suggests that the app needs to be regularly updated with new content to maintain its usefulness over time. These findings will be incorporated into the further development of the *Enable* app. We concluded from patients' feedback and requests that similar mHealth apps could benefit from giving patients a more active role (eg, being able to actively document side effects as they show up instead of being prompted to do so). In addition, regular updates to app content (eg, adding new informational pieces) could further contribute to and, thus, encourage the continued use of mHealth apps.

XSL•FO
RenderX

Acknowledgments

This study was funded by the Baden-Württemberg (Germany) Ministry of Science, Research, and Arts under reference 42-04HV.MED (19)/15/1 as part of the Center for Innovative Care project. The authors would like to thank all patients in this study for their participation and valuable input. They would also like to thank Corinna Hartmann for her support in recruiting patients for this study. Thank you to Bendix Harms for his support with the quantitative analysis.

Data Availability

The data sets used and analyzed during this study are available from the corresponding author upon reasonable request.

Authors' Contributions

CA, PM, and LW collaborated on the draft of the manuscript, PK contributed. MW is the principal investigator of the *ENABLE* project. LW, JM, and PM were responsible for the study design and protocol, and OH contributed. MW and TMD prepared and submitted the study protocol, and LW, JM, and PM contributed. LW, JM, and PM collaborated on the construction and testing of the interview guides and the quantitative data collection tools, and LS supported the finalization of these instruments. CA and LW conducted the interviews, analyzed transcripts, and interpreted the interview data. PM, CA, and LW conducted the eye-tracking study and retrospective interviews, TL contributed. PM analyzed and interpreted the eye-tracking data. PM analyzed and interpreted the survey data, and CA and LW supported data interpretation. OH contributed to the acquisition of funding. All the authors provided substantial comments and approved the final version of the manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Interview guide.
[DOCX File , 27 KB - humanfactors_v11i1e50926_app1.docx]

Multimedia Appendix 2
COREQ (Consolidated Criteria for Reporting Qualitative Research) checklist.
[PDF File (Adobe PDF File), 480 KB - humanfactors_v11i1e50926_app2.pdf]

Multimedia Appendix 3
Eye-tracking tasks.
[DOCX File , 18 KB - humanfactors_v11i1e50926_app3.docx]

Multimedia Appendix 4
Definition of themes.
[DOCX File , 34 KB - humanfactors_v11i1e50926_app4.docx]

Multimedia Appendix 5
Task performance data.
[XLSX File (Microsoft Excel File), 15 KB - humanfactors_v11i1e50926_app5.xlsx]

References

1. Erdmann F, Spix C, Katalinic A, Christ M, Folkerts J, Hansmann J, et al. Krebs in Deutschland für 2017/2018. Robert Koch Institut. 2021. URL: https://edoc.rki.de/handle/176904/9042 [accessed 2024-02-19]
2. Tighe M, Molassiotis A, Morris J, Richardson J. Coping, meaning and symptom experience: a narrative approach to the overwhelming impacts of breast cancer in the first year following diagnosis. Eur J Oncol Nurs 2011 Jul;15(3):226-232 [FREE Full text] [doi: 10.1016/j.ejon.2011.03.004] [Medline: 21511530]
3. ElKefi S, Asan O. How technology impacts communication between cancer patients and their health care providers: a systematic literature review. Int J Med Inform 2021 May;149:104430 [FREE Full text] [doi: 10.1016/j.ijmedinf.2021.104430] [Medline: 33684711]
4. Kashgary A, Alsolaimani R, Mosli M, Faraj S. The role of mobile devices in doctor-patient communication: a systematic review and meta-analysis. J Telemed Telecare 2017 Sep;23(8):693-700. [doi: 10.1177/1357633X16661604] [Medline: 27632990]
5. Qudah B, Luetsch K. The influence of mobile health applications on patient - healthcare provider relationships: a systematic, narrative review. Patient Educ Couns 2019 Jun;102(6):1080-1089. [doi: 10.1016/j.pec.2019.01.021] [Medline: 30745178]

XSL•FO
RenderX

6. Jongerius C, Russo S, Mazzocco K, Pravettoni G. Research-tested mobile apps for breast cancer care: systematic review. JMIR Mhealth Uhealth 2019 Feb 11;7(2):e10930 [FREE Full text] [doi: 10.2196/10930] [Medline: 30741644]
7. Chen L, Bartel C, Cai X, Cheng Y, Perer A, McClaine S, et al. Patient and provider perspectives on symptom monitoring during outpatient chemotherapy: interview study. JMIR Form Res 2023 Apr 17;7:e46001 [FREE Full text] [doi: 10.2196/46001] [Medline: 37067857]
8. Basch E, Deal AM, Kris MG, Scher HI, Hudis CA, Sabbatini P, et al. Symptom monitoring with patient-reported outcomes during routine cancer treatment: a randomized controlled trial. J Clin Oncol 2016 Feb 20;34(6):557-565 [FREE Full text] [doi: 10.1200/JCO.2015.63.0830] [Medline: 26644527]
9. Di Maio M, Gallo C, Leighl NB, Piccirillo MC, Daniele G, Nuzzo F, et al. Symptomatic toxicities experienced during anticancer treatment: agreement between patient and physician reporting in three randomized trials. J Clin Oncol 2015 Mar 10;33(8):910-915 [FREE Full text] [doi: 10.1200/JCO.2014.57.9334] [Medline: 25624439]
10. Zapata BC, Fernández-Alemán JL, Idri A, Toval A. Empirical studies on usability of mHealth apps: a systematic literature review. J Med Syst 2015 Feb;39(2):1. [doi: 10.1007/s10916-014-0182-2] [Medline: 25600193]
11. Or C, Tao D. Usability study of a computer-based self-management system for older adults with chronic diseases. JMIR Res Protoc 2012 Nov 08;1(2):e13 [FREE Full text] [doi: 10.2196/resprot.2184] [Medline: 23612015]
12. Nielsen J. What is usability? In: User Experience Re-Mastered: Your Guide to Getting the Right Design. Cambridge, MA: Morgan Kaufmann; 2010:3-22.
13. ISO 9241-11:2018 Ergonomics of human-system interaction part 11: usability: definitions and concepts. International Organization for Standardization. 2018 Mar. URL: https://www.iso.org/standard/63500.html [accessed 2024-02-19]
14. Law EL, Abrahão S. Interplay between user experience (UX) evaluation and system development. Int J Hum Comput Stud 2014 Jun;72(6):523-525. [doi: 10.1016/j.ijhcs.2014.03.003]
15. Mendes-Santos C, Nunes F, Weiderpass E, Santana R, Andersson G. Development and evaluation of the usefulness, usability, and feasibility of iNNOV breast cancer: mixed methods study. JMIR Cancer 2022 Feb 15;8(1):e33550 [FREE Full text] [doi: 10.2196/33550] [Medline: 35166682]
16. Cai T, Huang Y, Zhang Y, Lu Z, Huang Q, Yuan C. Mobile health applications for the care of patients with breast cancer: a scoping review. Int J Nurs Sci 2021 Aug 25;8(4):470-476 [FREE Full text] [doi: 10.1016/j.ijnss.2021.07.003] [Medline: 34631997]
17. Sadegh SS, Khakshour Saadat P, Sepehri MM, Assadi V. A framework for m-health service development and success evaluation. Int J Med Inform 2018 Apr;112:123-130. [doi: 10.1016/j.ijmedinf.2018.01.003] [Medline: 29500009]
18. Harte R, Glynn L, Rodríguez-Molinero A, Baker PM, Scharf T, Quinlan LR, et al. A human-centered design methodology to enhance the usability, human factors, and user experience of connected health systems: a three-phase methodology. JMIR Hum Factors 2017 Mar 16;4(1):e8 [FREE Full text] [doi: 10.2196/humanfactors.5443] [Medline: 28302594]
19. Doyle L, Brady AM, Byrne G. An overview of mixed methods research – revisited. J Res Nurs 2016;21(8):623-635. [doi: 10.1177/1744987116674257]
20. Bryman A. Integrating quantitative and qualitative research: how is it done? Qual Res 2006;6(1):97-113. [doi: 10.1177/1468794106058877]
21. Lim JY, Kim Y, Yeo SM, Chae BJ, Yu J, Hwang JH. Feasibility and usability of a personalized mHealth app for self-management in the first year following breast cancer surgery. Health Informatics J 2023;29(1):14604582231156476 [FREE Full text] [doi: 10.1177/14604582231156476] [Medline: 36772832]
22. Stubbins R, He T, Yu X, Puppala M, Ezeana CF, Chen S, et al. A behavior-modification, clinical-grade mobile application to improve breast cancer survivors' accountability and health outcomes. JCO Clin Cancer Inform 2018 Dec;2:1-11 [FREE Full text] [doi: 10.1200/CCI.18.00054] [Medline: 30652617]
23. Cooke L. Is eye tracking the next step in usability testing? In: Proceedings of the 2006 IEEE International Professional Communication Conference. 2006 Presented at: 2006 IEEE; October 23-25, 2006; Saragota Springs, NY. [doi: 10.1109/IPCC.2006.320355]
24. Jacob RJ, Karn KS. Commentary on section 4 - eye tracking in human-computer interaction and usability research: ready to deliver the promises. In: The Mind's Eye: Cognitive and Applied Aspects of Eye Movement Research. Amsterdam, The Netherlands: Elsevier Science; 2003:573-605.
25. Zhou L, Bao J, Setiawan IM, Saptono A, Parmanto B. The mHealth App Usability Questionnaire (MAUQ): development and validation study. JMIR Mhealth Uhealth 2019 Apr 11;7(4):e11500 [FREE Full text] [doi: 10.2196/11500] [Medline: 30973342]
26. Moorthy P, Weinert L, Harms BC, Anders C, Siegel F. German version of the mHealth app usability questionnaire in a cohort of patients with cancer: translation and validation study. JMIR Hum Factors 2023 Nov 01;10:e51090 [FREE Full text] [doi: 10.2196/51090] [Medline: 37910144]
27. Tong A, Sainsbury P, Craig J. Consolidated criteria for reporting qualitative research (COREQ): a 32-item checklist for interviews and focus groups. Int J Qual Health Care 2007 Dec;19(6):349-357. [doi: 10.1093/intqhc/mzm042] [Medline: 17872937]

XSL•FO
RenderX

28. Harris PA, Taylor R, Thielke R, Payne J, Gonzalez N, Conde JG. Research electronic data capture (REDCap) - A metadata-driven methodology and workflow process for providing translational research informatics support. J Biomed Inform 2009 Apr;42(2):377-381 [FREE Full text] [doi: 10.1016/j.jbi.2008.08.010] [Medline: 18929686]
29. Braun V, Clarke V. Using thematic analysis in psychology. Qual Res Psychol 2006;3(2):77-101. [doi: 10.1191/1478088706qp063oa]
30. Bansback N, Li LC, Lynd L, Bryan S. Development and preliminary user testing of the DCIDA (dynamic computer interactive decision application) for 'nudging' patients towards high quality decisions. BMC Med Inform Decis Mak 2014 Aug 01;14:62 [FREE Full text] [doi: 10.1186/1472-6947-14-62] [Medline: 25084808]
31. Erol Barkana D, Açık A, Duru DG, Duru AD. Improvement of design of a surgical interface using an eye tracking device. Theor Biol Med Model 2014 May 07;11 Suppl 1(Suppl 1):S4 [FREE Full text] [doi: 10.1186/1742-4682-11-S1-S4] [Medline: 25080176]
32. Rashid S, Soo ST, Sivaji A, Naeni HS, Bahri S. Preliminary usability testing with eye tracking and FCAT analysis on occupational safety and health websites. Procedia Soc Behav Sci 2013 Nov 06;97:737-744. [doi: 10.1016/j.sbspro.2013.10.295]
33. Eghdam A, Forsman J, Falkenhav M, Lind M, Koch S. Combining usability testing with eye-tracking technology: evaluation of a visualization support for antibiotic use in intensive care. Stud Health Technol Inform 2011;169:945-949. [Medline: 21893885]
34. Forsman J, Anani N, Eghdam A, Falkenhav M, Koch S. Integrated information visualization to support decision making for use of antibiotics in intensive care: design and usability evaluation. Inform Health Soc Care 2013 Dec;38(4):330-353. [doi: 10.3109/17538157.2013.812649] [Medline: 23957739]
35. Russell DM, Chi EH. Looking back: retrospective study methods for HCI. In: Olson J, Kellogg W, editors. Ways of Knowing in HCI. New York, NY: Springer; 2014.
36. Ansaar MZ, Hussain J, Bang J, Lee S, Shin KY, Woo KY. The mHealth applications usability evaluation review. In: Proceedings of the 2020 International Conference on Information Networking (ICOIN). 2020 Presented at: ICOIN; January 7-10, 2020; Barcelona, Spain. [doi: 10.1109/icoin48656.2020.9016509]
37. Kjeldskov J, Skov MB, Stage J. A longitudinal study of usability in health care: does time heal? Int J Med Inform 2010 Jun;79(6):e135-e143 [FREE Full text] [doi: 10.1016/j.ijmedinf.2008.07.008] [Medline: 18757234]
38. Dunn Lopez K, Chin CL, Leitão Azevedo RF, Kaushik V, Roy B, Schuh W, et al. Electronic health record usability and workload changes over time for provider and nursing staff following transition to new EHR. Appl Ergon 2021 May;93:103359. [doi: 10.1016/j.apergo.2021.103359] [Medline: 33556884]
39. Sonderegger A, Zbinden G, Uebelbacher A, Sauer J. The influence of product aesthetics and usability over the course of time: a longitudinal field experiment. Ergonomics 2012 Apr 16;55(7):713-730. [doi: 10.1080/00140139.2012.672658] [Medline: 22506866]
40. Nápoles AM, Santoyo-Olsson J, Chacón L, Stewart AL, Dixit N, Ortiz C. Feasibility of a mobile phone app and telephone coaching survivorship care planning program among Spanish-speaking breast cancer survivors. JMIR Cancer 2019 Jul 09;5(2):e13543 [FREE Full text] [doi: 10.2196/13543] [Medline: 31290395]
41. Petrocchi S, Filipponi C, Montagna G, Bonollo M, Pagani O, Meani F. A breast cancer smartphone app to navigate the breast cancer journey: mixed methods study. JMIR Form Res 2021 May 10;5(5):e28668 [FREE Full text] [doi: 10.2196/28668] [Medline: 33970120]
42. Wildenbos GA, Jaspers MW, Schijven MP, Dusseljee-Peute LW. Mobile health for older adult patients: using an aging barriers framework to classify usability problems. Int J Med Inform 2019 Apr;124:68-77. [doi: 10.1016/j.ijmedinf.2019.01.006] [Medline: 30784429]
43. Morey SA, Barg-Walkow LH, Rogers WA. Managing heart failure on the go: usability issues with mHealth apps for older adults. Proc Hum Factors Ergon Soc Annu Meet 2017 Sep 28;61(1):1-5. [doi: 10.1177/1541931213601496]
44. Durchschnittsalter der bevölkerung in Deutschland (1871-2021). Bundesinstitut für Bevölkerungsforschung. URL: https://www.bib.bund.de/DE/Fakten/Fakt/B19-Durchschnittsalter-Bevoelkerung-ab-1871.html [accessed 2024-02-19]
45. Jessen S, Mirkovic J, Halvorsen Brendmo E, Solberg Nes L. Evaluating a strengths-based mHealth tool (MyStrengths): explorative feasibility trial. JMIR Form Res 2021 Nov 17;5(11):e30572 [FREE Full text] [doi: 10.2196/30572] [Medline: 34787580]
46. Graham TA, Ali S, Avdagovska M, Ballermann M. Effects of a web-based patient portal on patient satisfaction and missed appointment rates: survey study. J Med Internet Res 2020 May 19;22(5):e17955 [FREE Full text] [doi: 10.2196/17955] [Medline: 32427109]
47. Avdagovska M, Ballermann M, Olson K, Graham T, Menon D, Stafinski T. Patient portal implementation and uptake: qualitative comparative case study. J Med Internet Res 2020 Jul 27;22(7):e18973 [FREE Full text] [doi: 10.2196/18973] [Medline: 32716308]
48. Müller J, Weinert L, Svensson L, Rivinius R, Kreusser MM, Heinze O. Mobile access to medical records in heart transplantation aftercare: mixed-methods study assessing usability, feasibility and effects of a mobile application. Life (Basel) 2022 Aug 08;12(8):1204 [FREE Full text] [doi: 10.3390/life12081204] [Medline: 36013383]

XSL•FO
RenderX

49. Niazkhani Z, Toni E, Cheshmekaboodi M, Georgiou A, Pirnejad H. Barriers to patient, provider, and caregiver adoption and use of electronic personal health records in chronic care: a systematic review. BMC Med Inform Decis Mak 2020 Jul 08;20(1):153 [FREE Full text] [doi: 10.1186/s12911-020-01159-1] [Medline: 32641128]
50. Sollie M, Hansen M, Thomsen JB. Health technology readiness amongst patients with suspected breast cancer using the READHY-tool - a cross-sectional study. J Med Syst 2023 Nov 16;47(1):118 [FREE Full text] [doi: 10.1007/s10916-023-02016-0] [Medline: 37971517]

Abbreviations

COREQ: Consolidated Criteria for Reporting Qualitative Research
ISO: International Organization for Standardization
MAUQ: mHealth App Usability Questionnaire
mHealth: mobile health
PRO: patient-reported outcome
QoL: quality of life
RCT: randomized controlled trial
REDCap: Research Electronic Data Capture

Edited by P Santana-Mancilla; submitted 18.07.23; peer-reviewed by A Carrillo, M Moreno, A Wada, C Baxter; comments to author 25.08.23; revised version received 30.11.23; accepted 17.12.23; published 05.03.24.

Please cite as:
Anders C, Moorthy P, Svensson L, Müller J, Heinze O, Knaup P, Wallwiener M, Deutsch TM, Le TV, Weinert L
Usability and User Experience of an mHealth App for Therapy Support of Patients With Breast Cancer: Mixed Methods Study Using Eye Tracking
JMIR Hum Factors 2024;11:e50926
URL: https://humanfactors.jmir.org/2024/1/e50926
doi:10.2196/50926
PMID:38441959

©Carolin Anders, Preetha Moorthy, Laura Svensson, Julia Müller, Oliver Heinze, Petra Knaup, Markus Wallwiener, Thomas M Deutsch, Thao-Vy Le, Lina Weinert. Originally published in JMIR Human Factors (https://humanfactors.jmir.org), 05.03.2024. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on https://humanfactors.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Feasibility and Design Factors for Home-Based Pulmonary Rehabilitation of Patients With Chronic Obstructive Pulmonary Disease and Chronic Lung Diseases Based on a People-Object-Environment Framework: Qualitative Interview Study

Shih-Ying Chien[1,2], PhD; Alice May-Kuen Wong[2], MD; Winston Tseng[3], PhD; Han-Chung Hu[4,5], MD; Hsiu-Ying Cho[5], MA

[1]Department of Industrial Design, Chang Gung University, Taoyuan, Taiwan
[2]Department of Physical Medicine and Rehabilitation, Chang Gung Memorial Hospital, Taoyuan, Taiwan
[3]Division of Community Health Sciences, School of Public Health, University of California, Berkeley, CA, United States
[4]Department of Thoracic Medicine, Chang Gung Memorial Hospital, Taoyuan, Taiwan
[5]Department of Respiratory Therapy, Lin-Kou Chang Gung Memorial Hospital, Taoyuan, Taiwan

Corresponding Author:
Shih-Ying Chien, PhD
Department of Industrial Design
Chang Gung University
259 Wen-Hwa 1st Rd, Kweishan
Taoyuan, 333
Taiwan
Phone: 886 3 2118800 ext 5669
Email: shihying@mail.cgu.edu.tw

Abstract

Background: The feasibility of implementing home-based pulmonary rehabilitation (PR) can be assessed from the perspectives of patients with chronic lung disease and health care professionals involved in PR.

Objective: Through a qualitative inquiry using interviews and the adoption of the people-object-environment framework, this study aims to understand the influences of interpersonal, environmental, and situational factors on the perceptions and considerations of individuals involved in home-based PR for patients with chronic lung disease.

Methods: One-on-one interviews were conducted with 20 patients with chronic lung disease and 20 health care professionals for investigating their attitudes and opinions based on their experiences regarding home-based PR as well as for identifying the key factors affecting the benefits and drawbacks of such therapies. This study further evaluates the feasibility of using digital tools for medical diagnosis and treatment by examining the technology usage of both parties.

Results: The 4 key issues that all participants were the most concerned about were as follows: distance to outpatient medical care, medical efficiency, internet connectivity and equipment, and physical space for diagnosis and treatment. Interviews with patients and health care professionals revealed that the use of technology and internet was perceived differently depending on age and area of residence. Most participants reported that digital tools and internet connectivity had many benefits but still could not solve all the problems; moreover, these same digital tools and network transmission could lead to problems such as information security and digital divide concerns. This study also emphasizes the significant impact of human behavior and thinking on shaping the design of health care interventions and technologies. Understanding user perspectives and experiences is crucial for developing effective solutions for unmet needs.

Conclusions: The results of this study indicate that despite the different perspectives of patients and health care professionals, their considerations of the key issues are very similar. Therefore, the implementation of plans related to telemedicine diagnosis, treatment, or rehabilitation should take the suggestions and considerations of both parties into account as crucial factors for telehealth care design.

XSL•FO
RenderX

(JMIR Hum Factors 2024;11:e51150) doi:10.2196/51150

KEYWORDS

chronic lung diseases; home-based pulmonary rehabilitation; telehealth; remote health care

Introduction

As the third leading cause of morbidity and mortality worldwide, chronic obstructive pulmonary disease (COPD) is a significant public health issue [1-4]. In 2019, the number of individuals diagnosed with COPD exceeded 328 million worldwide [5-8]. A significant correlation between physical activity and lung function [9-12] emphasizes the importance of regular exercise for individuals with COPD who require pulmonary rehabilitation (PR) [13-19]. However, patients with COPD often report reluctance to engage in physical activities due to dyspnea, the effects of which include chronic cough, exacerbations, reduced exercise capacity, and impaired quality of life [20-25]. PR is a tailored and comprehensive intervention conducted via a thorough assessment of the patient. In individuals with chronic pulmonary diseases, the primary objective of the pulmonary intervention is to improve not only their overall health but also their psychosocial well-being in the long term [26-28]. Typically, PR programs are customized for personal symptomatic conditions [29-31]; hence, PR interventions entail tailored exercises and educational sessions aimed at enhancing activity tolerance, mitigating symptoms, and augmenting skills that aid in managing chronic respiratory diseases [31,32]. The majority of PR treatments usually require one-on-one sessions and the assistance of a therapist [33-35]. However, the one-on-one care approach is limited due to shortages in health care personnel, elevated work-related stress, and prolonged working hours [36]. Moreover, when the COVID-19 pandemic hit, lockdowns and personnel restrictions forced the interruption of PR for many patients with chronic lung disease, which posed a threat to their lives [37-40].

Due to the COVID-19 pandemic, telehealth has become increasingly attractive owing to its functionality, importance, and prospects [41,42]. In addition to reducing human contact and easing the burden on health care workers, telehealth leverages technology communication and transmission to alleviate the workload of respiratory therapists and improve the accuracy of respiratory rehabilitation records [43-45]. Using telehealth, patients can undergo rehabilitation at home and be monitored remotely by medical personnel [46,47]. Home-based PR can also mitigate the difficulties of outpatient care for patients living in remote areas and those with physical disabilities [48-50]. Furthermore, it can be used as an auxiliary means of physical PR to assist in self-management and precisely modify behavior, thereby reducing hospitalization and medical costs [51,52].

Traditional PR usually relies on one-on-one human monitoring through observation or physiological monitors to examine a patient's health condition. Remote health care has the advantage of prescribing home-based PR, enabling patients who are unable to leave their homes due to physical conditions such as disability or living in rural areas to partake in rehabilitation programs at home [53,54]. However, there are also many limitations and considerations of remote health care, as follows:

1. Lack of security and limited interpersonal interaction: The safety of patients is the primary concern of clinical physicians [55,56]. The biggest challenge of home rehabilitation is emergency treatment, which has been the main hurdle for remote health care since many years [57]. In addition, remote therapy can only provide limited physical and mental assessments [58,59]. Due to the lack of face-to-face interpersonal interactions, patients may develop loneliness, helplessness, and frustration, which may reduce the effectiveness of treatments and the speed of recovery [60,61].
2. Privacy and security issues: Most remote health care is performed through network transmission. Many clinical physicians believe that network transmission may lead to data leakage or theft of medical records or personal information of patients [62,63].
3. Technological and equipment limitations: The implementation of remote health care requires specific technological equipment such as smartphones or computers with network functions. However, for many remote users or special groups such as older persons, lack of equipment, poor network communication quality, or unfamiliarity with network-related technology hinder utilization [64].
4. Insurance payment limitations: Different regions or countries have different standards for remote health care services. Therefore, many insurance companies do not have a remote health care reimbursement system or only cover specific services [65-68].

Despite its limitations and by taking people, object, and environment into consideration, telemedicine remains a valuable tool for the provision of health care services, especially for patients who have difficulty visiting medical facilities in person or those affected by infectious diseases and related restrictions such as lockdowns and quarantine. Telemedicine enables uninterrupted treatment and continued assistance for patients in their recovery. However, in establishing a home-based PR, it is essential to consider the various environments of participants to effectively maximize the benefits of this medical service.

Methods

Ethics Approval

This study was approved by the institutional review board of Chang Gung Memorial Hospital (approval 202200070B0). The participants were patients with chronic lung disease and respiratory health care professionals who had provided written informed consent from both urban and rural areas. Due to the COVID-19 pandemic, all one-on-one interviews were conducted by videoconferencing.

XSL•FO
RenderX

Participants and Procedures

The 20 patients recruited for the interviews included those who had participated in PR programs and those who had not. During the interviews, the patients provided insights into the implementation of PR programs from a patient-centric standpoint. All interviewees had a medical history of 5 years or more.

The 20 respiratory health care professionals included registered thoracic surgeons, respiratory therapists, physical therapists, and PR specialists. Most of these professionals had experience treating patients with chronic lung disease and had participated in designing exercise prescriptions, patient tracking and monitoring, and disease progression research in PR programs. Furthermore, the majority of the interviewees had treated a specific proportion of patients with chronic lung disease within the past 3 years (Table 1).

Table 1. Characteristics of the health care professionals (n=20).

	Values
Gender (male:female)	8:12
Age (years), min-max; mean (SD)	27-65; 46 (11)
Experience in pulmonary rehabilitation (years), mean (SD)	5.2 (8.47)
Type of health care professional in pulmonary rehabilitation, n (%)	
Thoracic surgeons	3 (15)
Respiratory therapists	12 (60)
Physical therapists	2 (10)
Pulmonary rehabilitation specialists	3 (15)

Prior to the interviews, all participants were required to complete a survey questionnaire, which included demographic information and details of their use of smart devices and the internet. Daily use was defined as regular usage. The patients provided information about their pulmonary disease status, duration of illness, and a self-assessment of their health status (on a 5-point scale ranging from excellent to poor) as well as recalled their activity frequency over the past 7 days. Health care personnel were required to answer questions related to their primary clinical responsibilities. Each participant took part in a 1.5- to 2-hour interview session conducted by the primary author, who was also a clinical researcher and an assistant professor affiliated with the Chang Gung Medical Foundation. In-depth interviews were primarily used to collect the data. After collecting the interview data, all identifiable personal information was removed from the transcripts. The data were then coded, organized, and analyzed using NVivo 12.0 software (Lumivero) for qualitative data analysis. For accurate and detailed data interpretation, the transcripts were provided to the interviewees for review and cross-checked with relevant researchers to confirm the accuracy of data interpretation.

Results

Characteristics of the Participants

This study consisted of 40 participants: 20 health care professionals specializing in PR and 20 patients with chronic lung diseases. The background characteristics of the 20 health care professionals are shown in Table 1; nearly 60% (12/20) were respiratory therapists, and the remaining health care professionals were pulmonary surgeons, physical therapists, and rehabilitation physicians. Their mean age was 46 years, and all had more than 3 years of experience in PR and treatment (mean 5.2 years). The background characteristics and activity habits of the 20 patients interviewed are shown in Table 2; the majority of the patients had COPD (12/20, 60%), and 25% (5/20) were lung transplant recipients. The majority of the participants (15/20, 74%) had never participated in a PR program, and 70% (14/20) of the patients rated their physical condition as poor. Regarding exercise over the past week, 65% (13/20) of the patients chose a 10-minute walk as their exercise indicator, followed by strength training (5/20, 25%). Notably, 55% (11/20) of the patients reported preferring to sit rather than stand and to stand rather than move.

XSL•FO
RenderX

Table 2. Characteristics of the patients (n=20).

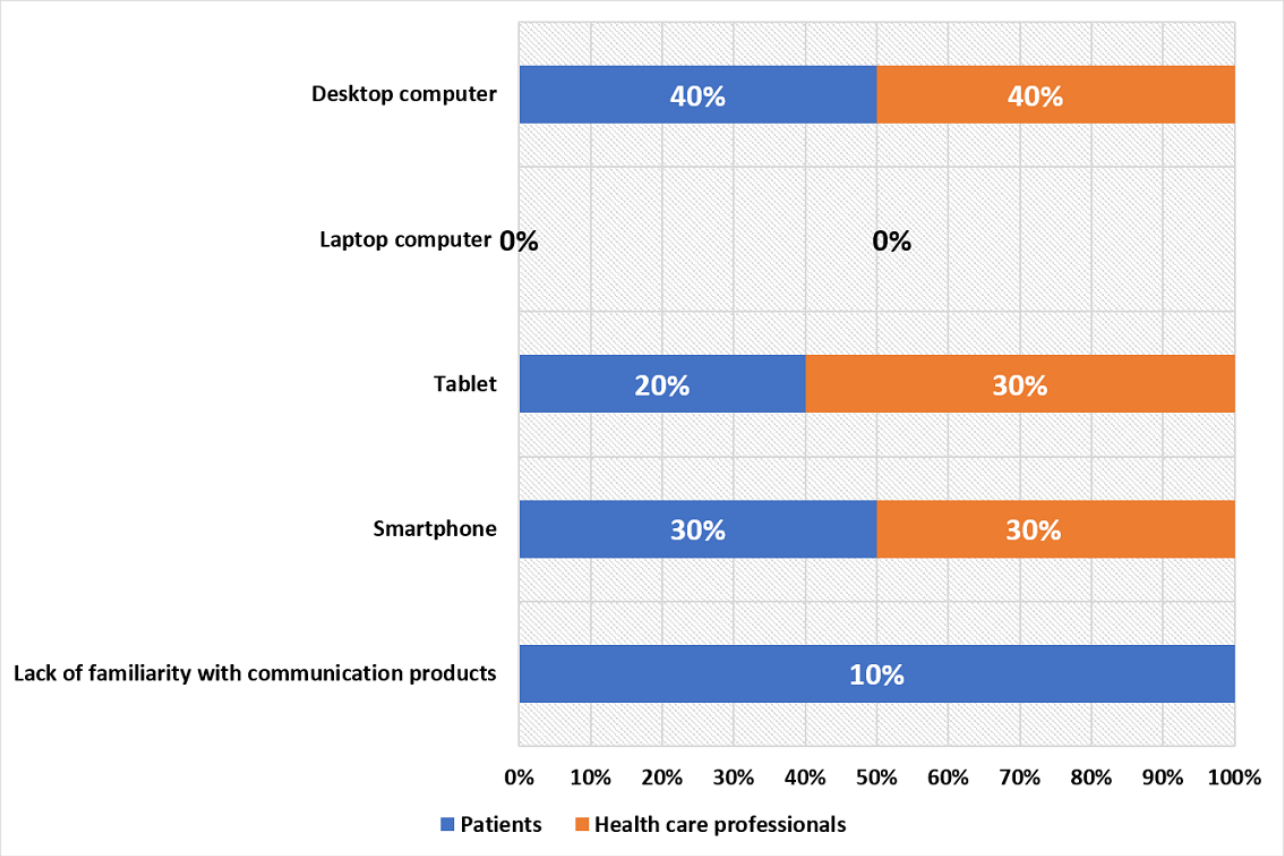
	Values
Gender (male:female)	10:10
Age (years), min-max; mean (SD)	51-85; 68 (9.8)
Participation in pulmonary rehabilitation programs, n (%)	
Yes	5 (26)
No	15 (74)
Chronic lung diseases, n (%)	
Chronic obstructive pulmonary disease	12 (60)
Asthma	3 (15)
Lung transplantation	5 (25)
Self-assessment of their health, n (%)	
Excellent	0 (0)
Very good	0 (0)
Good	1 (5)
Fair	5 (25)
Poor	14 (70)
Exercise frequency and quantity over the past week, n (%)	
I have engaged in high-intensity strength training, including aerobic exercise, fast cycling, and swimming.	2 (10)
I have participated in moderate physical activities such as stretching exercises and flexibility training.	5 (25)
I have walked for at least 10 minutes every day.	13 (65)
What statement best characterizes my exercise habits? n (%)	
Given the option, I will opt to sit rather than stand.	11 (55)
I frequently require standing but not for the purpose of lifting heavy objects.	5 (25)
Climbing slopes and stairs is a common necessity for me.	4 (20)
I often transport heavy objects and engage in manual labor.	0 (0)

Survey Results

In order to better examine the potential of telehealth, we conducted a survey targeting contemporary electronic communication tools, specifically computers, cellphones, and tablets, which were widely utilized by both patients and health care professionals, as shown in Figure 1. The majority of the patients and health care professionals reported using desktop computers as their most frequently used electronic communication device, constituting the largest proportion at 40% (8/20), followed by smartphones at 30% (6/20). Notably, health care professionals reported a higher usage rate (by 10%) of tablet computers compared to patients. Note that neither group of participants reported habitually using laptops. Overall, the patients were less proficient with technology compared to the health care professionals, which is a crucial determinant in implementing telehealth programs.

XSL•FO
RenderX

Figure 1. Survey results of the usage of electronic communication devices among patients and health care professionals in this study.



People-Object-Environment Framework

This study adopts the people-object-environment framework as the focal point for the interview investigation to improve the understanding of the feasibility of home-based PR. This study analyzes the advantages and disadvantages of remote PR in the current context, with the aim of bridging the gap between ideal use and reality (Table 3). Activities involving various elements such as individuals, entities, and environmental factors often result in the emergence of diverse concerns among different participants. Through the analysis presented in Table 3, we identified the gaps in home-based rehabilitation services from the perspectives of health care professionals and patients. Subsequently, this facilitated a thorough discussion of potential solutions to meet the needs and expectations of all involved parties. Our research findings reveal that the use of telehealth for home-based PR programs had both advantages and disadvantages. Using a people-object-environment framework to analyze the results, we describe 4 dimensions: reduced time and transportation constraints to access medical care; improved medical efficiency; changes in equipment, network, and physical space; and information transmission security, about which health care professionals particularly raised concerns in telehealth.

XSL•FO
RenderX

Table 3. Analysis of the pros and cons of telerehabilitation with the people-object-environment framework.

	People (patient)	Object (patient)	Environment (patient)
People (health care professionals)			
Pros	• Promoting health care access in remote areas: Telehealth facilitates convenient health care services, which enhance medical care for patients in remote areas and promote community health. • Boosting patient involvement: Telehealth enables interactions with health care professionals via web-based platforms, providing medical information and guidance as well as fostering active engagement of patients in their own health management.	• Improving health care resource allocation: Telehealth enables physicians to diagnose and treat patients across different geographical areas, alleviating shortages in local health care resources and enhancing the efficiency of health care resource allocation.	• Reducing health care burden: Telehealth reduces the health care burden for long-term patients or those requiring regular follow-ups; this minimizes the time and effort associated with transportation and waiting as well as provides cost-effective health care options. • Decreasing cross-infection risks: Telehealth minimizes contact between patients and health care professionals, thereby lowering the risk of cross-infection and promoting the health and safety of both health care professionals and patients.
Cons	• Bridging communication barriers: Telehealth reduces physical interactions and social contact between patients and physicians, which may have long-term effects on patients' psychological and social well-being.	• Operational and communication barriers: Older or technologically inexperienced patients may encounter difficulties in understanding instructions from remote health care professionals through telehealth.	• Environmental limitations: Home environments often impose spatial constraints that may limit various rehabilitation, diagnostic, and treatment activities.
Object (health care professionals)			
Pros	• Advantages of telehealth: Telehealth provides a convenient health care model, particularly beneficial for regions facing constraints related to time, geographical location, and transportation. The utilization of basic computer equipment enables the provision of medical consultations, making health care more accessible and efficient for patients in such areas.	• Wireless transmission: Wireless transmission significantly reduces the workload of health care professionals and makes health care services more efficient by transitioning from a one-on-one service to a one-to-many format.	• Digital health care infrastructure: Telehealth accelerates the transmission and exchange speed of health care information, leading to improved overall health care efficiency.
Cons	• Digital divide: Individuals who are from lower socioeconomic backgrounds or have limited access to digital resources may face barriers to participation in telehealth due to the lack of appropriate technological equipment or internet connectivity. This highlights inequalities in the distribution of health care resources. • Technological dependency: Users with limited technological skills or resources may encounter difficulties in operating telehealth, which requires adequate knowledge of technology, suitable equipment, and stable internet connectivity. • Health care quality and patient experience: Although telehealth provides convenient remote health care options for certain diseases or conditions, in-person consultations or measurements from medical instruments may offer more accurate health care services.	• Disparity in health care resources between urban and rural areas: Despite the convenience of telehealth for remote consultations, operational difficulties may still exist for areas lacking proper equipment.	• Equipment and infrastructure requirements: Telehealth relies on high-speed internet and appropriate equipment, which can still pose challenges in certain rural areas.
Environment (health care professionals)			

XSL•FO
RenderX

	People (patient)	Object (patient)	Environment (patient)
Pros	• Expansion of health care service areas: Through telehealth, physicians can diagnose and treat patients remotely without being limited by geographical location while providing real-time medical services. • Expansion of professional scope: Telehealth enables physicians to engage in remote meetings and collaborations with other health care experts, enhancing medical efficiency. • Enhancement of diagnosis and treatment efficiency: Telehealth reduces time and space limitations between physicians and patients, improving the efficiency of the overall health care services. • Increased convenience: Telehealth offers patients greater convenience, particularly for those residing in remote areas or facing mobility challenges, thereby reducing the time and costs associated with hospital visits.	• Enhancement of diagnostic and treatment capabilities: Through remote imaging and information sharing, physicians can access additional support and assistance, which improve diagnostic accuracy and treatment outcomes. • Improvement of health care resource utilization: Telehealth aids physicians in managing and allocating regional health care resources more effectively, enhancing utilization efficiency and reducing unnecessary health care costs.	• Reducing reliance on physical space: Telehealth reduces the need for physical space such as clinics and hospitals, thereby lowering costs and burdens associated with facilities and resources for health care institutions.
Cons	• Lack of physical contact: Telehealth may not provide opportunities for face-to-face contact with patients, which can make it challenging for physicians to conduct comprehensive physical examinations or assessments. • Limitations in comprehensive treatment: Some diagnoses and treatments may require physical contact and assistance from specific equipment, which cannot necessarily be substituted by telehealth.	• Technical requirements: The use of telehealth requires stable internet connectivity and appropriate device support, which may be challenging for users who are not familiar with technology. • Medical responsibility and risk management: Telehealth may involve issues of medical responsibility and risk management, such as misdiagnosis, treatment errors, or incomplete medical records, which may result in medical disputes and litigation. Physicians and health care institutions need to ensure compliance with relevant medical responsibility and risk management principles in telehealth as well as maintain a high level of medical practice.	• Health care security and privacy risks: The use of telehealth may involve health care security and digital privacy risks such as patient identity verification and medical record protection. Physicians should exercise caution in handling such issues.

Dimensions

Dimension 1: Distance, Time, and Transportation Issues

Both patients and health care professionals acknowledged the significant benefits of telehealth in addressing the challenges of distance in accessing medical care. Reductions in travel and wait times due to telehealth allowed patients to actively participate in their health care decision-making through web-based platforms. The digitization of medical records for better disease management was also facilitated. For residents in remote areas, telehealth eliminated geographical barriers to health care, promoted more efficient allocation of medical resources, and prevented the closure of regional hospitals and the “medical deserts” phenomenon. Nevertheless, health care professionals identified potential risks and quality-of-care issues associated with telehealth. Due to the absence of physical interactions, physicians were limited to relying on surface-level symptoms for diagnosis. Comprehensive physiological examinations and physical evaluations were also limited due to the lack of suitable equipment, which could in fact jeopardize the safety of critically ill patients who require urgent or emergency treatment. Moreover, prolonged social isolation resulting from telehealth may have adverse effects on patients’ psychological and social well-being.

Dimension 2: Enhancing Medical Efficiency Through the Use of Telehealth

Health care professionals described that the popularization of telehealth was due to its advantages in improving the efficiency of disseminating medical information. Through online platforms, health care professionals can access patients’ medical history instantaneously and collaborate to provide optimized treatment. This approach reduced constraints on patients’ time and space, expanded the scope of medical services, reduced the workload of health care professionals, and transformed the traditional one-on-one respiratory treatment mode into a one-to-many model. Health care professionals also noted that the advantages

XSL•FO
RenderX

of data and imaging arising from telehealth actually improved diagnostic and treatment efficiency, which could as a result achieve precision medicine. However, health care professionals also raised concerns about telehealth. The mode of transmitting medical information through data still harbored many risks and considerations such as diagnostic and treatment errors. Additionally, injuries (such as falls or respiratory distress) that could occur during treatment raised issues related to medical responsibility and risk management. Therefore, handling patient identity verification and medical records with caution when administering telehealth is crucial in order to safeguard patient privacy during medical treatments.

Dimension 3: Leveraging Internet Connectivity and Device-Based Solutions for Rehabilitation and Health Monitoring

Patients and health care professionals both identified that digital health care had the potential to significantly reduce medical wait times and enhance efficiency, which complemented the services of regional hospitals, lowered the medical burden of chronic patients, and eliminated limitations due to transportation, geography, and time. However, many challenges still remain with the use of digital tools for rehabilitation and monitoring systems. Most patients who require PR are older, aged ≥65 years, and unfamiliar with digital devices and networks, and they often lack the knowledge and understanding of how to install and configure such devices and applications. In addition, remote areas lack stable networks, technology, and equipment. This digital divide inhibits a subset of patients in certain areas from fully utilizing relevant medical services, highlighting the problem of an uneven distribution of medical resources.

Dimension 4: Advantages and Disadvantages of Converting Medical Spaces

The changing medical environment has brought many advantages to patients and health care institutions through telehealth, particularly during the COVID-19 pandemic, by reducing hospital-acquired infections and patients' reliance on medical space and resources, thereby alleviating the burden of health care costs. However, as previously mentioned, some health care professionals reported that not all diagnoses and treatments could be properly conducted remotely due to the availability or operation of equipment and the limitation of patients' home space and environment, which restrict the implementation of many treatment regimens. All the 4 dimensions supported by quotes from patients and health care professionals are shown in Table 4.

XSL•FO
RenderX

Table 4. Verbatim quotes supporting the main dimensions by patients and health care professionals.

Dimensions	Patient	Health care professional
Dimension 1: Distance to outpatient care		
Positive	*…During the pandemic, being able to have online consultations reduced a lot of my stress. I heard many of my friends around me were infected while at hospital.* [Female, 48 years old]	*…Telemedicine has reduced a lot of transport-related issues. Through online connections, we can access all of the patient's data and make more accurate assessments.* [HCP[a] #12]
Negative	*…I live in a very rural area where there are no taxis, so every time I see a doctor, I have to take four different buses. The journey alone takes me over three hours, so I avoid seeing a doctor if I can help it.* [Female, 64 years old]	*…To be honest, not every patient is suitable for telemedicine. For example, older patients may have difficulty understanding what I ask them to do. Also, some patients' conditions cannot be determined solely by questioning and require examination using medical instruments and devices, so it is difficult for me to make a diagnosis without a proper examination.* [HCP #8]
Dimension 2: Medical efficiency		
Positive	*…Because I get breathless when I walk, I try not to go out if I don't have to. I'm also afraid of falling when I go out, and I don't want to bother my children. So if I can see a doctor through a computer, I prefer that.* [Male, 66 years old]	*…The biggest advantage of telehealth is saving a lot of time and manpower. Of course, this refers to medical work that is more repetitive and lower risk. But I hope that in the future, online systems will have warning functions that can quickly let me know which patient has an issue that needs special attention.* [HCP #3]
Negative	*…I would rather see a real doctor. Just talking on the phone doesn't give me a feeling that I've really seen a doctor.* [Female, 72 years old]	*…To be honest, although the internet is convenient, I feel that its effectiveness is sometimes limited. Perhaps respiratory therapy needs to be divided into stages, and not every stage is suitable for being done at home. It may need to be classified/graded.* [HCP #17]
Dimension 3: Internet connectivity and equipment		
Positive	*…Every time I go out to see a doctor, I'm always in a rush and get so nervous that I forget to ask the doctor any questions. By seeing the doctor through a computer, I have more time to chat with the doctor.* [Female, 70 years old]	*…The internet connection is very convenient. As long as the health insurance card is inserted, all the patient's information can be accessed. Telehealth has not only changed a patient's medical treatment mode but also prevented many regional hospitals from closing down.* [HCP #11]
Negative	*…To be honest, I don't really understand the internet. If no one helps me set it up, I won't know how to see a doctor online. And if the doctor doesn't see me, how will they know what's wrong with me?* [Male, 76 years old]	*…Sometimes, the reason why I cannot wait for a patient is because the foreign caregiver has not set up the computer properly. When communicating with the patient through the computer, sometimes the elderly cannot understand, and it is also difficult to communicate with the caregiver. If I were there in person, I could still teach them how to do it.* [HCP #8]
Dimension 4: Space for diagnosis and treatment		
Positive	*…I am old and unable to move around easily. It would be best for me to see the doctor at home.* [Male, 86 years old]	*…A patient receiving online medical care at home will require much less space for us, such as waiting rooms and registration areas. It will also significantly reduce the demand and burden on staff.* [HCP #1]
Negative	*…There are many things that I cannot do at home. I need to have my blood pressure measured, but there is no one to help me at home. Also, I like to chat with people, but at home, there's only me.* [Female, 70 years old]	*…Online medical care now is quite good, with many complete functions such as registration, appointment progress, and electronic medical records. However, I personally have reservations about having many medical records stored in the cloud, as there is no absolute security. Also, if a patient falls at home, how to allocate responsibility and the risks involved are also concerns.* [HCP #14]

[a]HCP: health care professional.

Discussion

This study explores the perspectives and barriers of respiratory health care professionals and patients toward telehealth for rehabilitation for respiratory diseases. Based on the participant interviews, the use of home-based telerehabilitation for patients with lung diseases was perceived to have both advantages and disadvantages, which could be categorized under 4 domains: location, digital technology, internet connectivity, and physical space requirements. Unlike previous research [69], we adopted the people-object-environment framework and interviewed patients as well as health care professionals, with the aim of obtaining feedback from all participants in the same context. The 4 aspects mentioned above were found to be the most important concerns for health care institutions and patients.

XSL•FO
RenderX

Previous studies have reported that distance to outpatient care has a profound impact on chronic patients; in other words, the farther away from home, the lower is a patient's willingness to seek medical care [70]. However, Bhatt et al [71,72] highlighted that patients often exhibit a reluctance to engage in PR, regardless of proximity, for several reasons. Both groups of interviewees in our study reported that the provision of alternative options would reduce the number of hospital visits, which would benefit patients and health care institutions. Although digital health care has its limitations, leveraging the internet to expand regional hospital services is not only beneficial to the public but also makes medical services more effective [73]. In special circumstances such as the outbreak of a pandemic, issues such as patients being unable to attend in-person treatments cannot be ignored. It is undeniable that digital health care, in particular, spawns numerous benefits in such situations. However, there are still limitations to digital health care for people (doctors or the public), equipment and network, and the environment (urban or rural). For example, most older patients feel that only consultations in person with doctors generate the feeling of being treated. Furthermore, without physical examinations, it could be difficult for doctors to diagnose the cause of symptoms. Nonetheless, digital technology remains a good choice for patients with respiratory disease who do not want to venture outside or exercise; however, not every patient's home is equipped with remote medical devices or equipment, especially in rural areas [74]. Most patients with respiratory disease are also older, and without caregiver assistance, operating such devices can be difficult. In addition, due to limited professional knowledge, equipment, network, and living space, home care cannot replace all hospital diagnoses and rehabilitation. Therefore, in consideration of the findings from our study and a previous study [75], the implementation of hierarchical medical care requires that patients first undergo video consultations and then be referred to nearby medical institutions for appropriate treatment based on the severity of their condition. Patients can be referred to a larger medical center when necessary for treatment through an electronic referral platform between institutions. This approach not only effectively improves the utilization efficiency of medical resources but also significantly reduces medical expenses and transportation costs for patients.

This study has some limitations. First, as our study was conducted in a specific health care institution, our findings may not be generalizable to other regions or institutions. Second, most patients had poor health conditions and no prior experience with PR; thus, they relied only on limited experience and information. Lastly, the majority of the participants were older, which may have influenced their responses regarding computer use or internet issues. Additionally, health care service needs likely vary between urban and rural areas, and our study does not distinguish between the challenges and differences in home-based PR between these 2 types of areas. Future research should consider this aspect in their study design.

Most study participants reported that telehealth could greatly benefit patients with chronic pulmonary diseases; however, these benefits were not without limitations. Reflections on these limitations by patients or health care professionals revealed that telehealth is not suitable for all patients. For example, diagnosis and treatment via telehealth can only accomplish certain tasks and merely serve as a tool for preliminary diagnostic assessments. Nonetheless, preliminary assessments can determine whether a referral to a regional hospital or a large teaching hospital is necessary. This classification and referral system will also be applicable to rehabilitation therapy. Not all patients are suitable for home-based PR, considering patient safety, the required space and equipment, or the need for further precision testing, among other factors. In addition, although telehealth brought many conveniences to patients and health care professionals, both parties still faced significant psychological pressure. Patients noted that digital medicine lacked warmth, and they tended to prefer human care, while physicians had doubts about medical decision-making without the ability to perform physical examinations. The degree of control over digital technology was also an issue. Both parties lacked confidence that effective treatment could be achieved solely through the internet. Even though digital care has the advantage of long-term patient monitoring, some patients were unfamiliar with internet devices, and health care professionals were concerned that patients may not always respond correctly to instructions. Moreover, patients often neglect their physician's advice (such as following prescribed exercise schedules) due to lack of motivation and the need to physically meet with the physician. Both groups of study participants indicated that significant improvements in telerehabilitation technology were still needed, particularly for patients in rural areas or those who were older and living alone, who require more support and services.

Acknowledgments

This research was supported by the National Science and Technology Council (grant 111-2222-E-182-002-MY2). We also thank all the interviewees who provided insight and expertise that greatly assisted this research.

Conflicts of Interest

None declared.

References

1. Viegi G, Maio S, Pistelli F, Baldacci S, Carrozzi L. Epidemiology of chronic obstructive pulmonary disease: health effects of air pollution. Respirology 2006 Sep;11(5):523-532. [doi: 10.1111/j.1440-1843.2006.00886.x] [Medline: 16916323]

XSL•FO
RenderX

2. Cukic V, Lovre V, Dragisic D, Ustamujic A. Asthma and chronic obstructive pulmonary disease (COPD) - differences and similarities. Mater Sociomed 2012;24(2):100-105 [FREE Full text] [doi: 10.5455/msm.2012.24.100-105] [Medline: 23678316]
3. Chapman KR, Mannino DM, Soriano JB, Vermeire PA, Buist AS, Thun MJ, et al. Epidemiology and costs of chronic obstructive pulmonary disease. Eur Respir J 2006 Jan;27(1):188-207 [FREE Full text] [doi: 10.1183/09031936.06.00024505] [Medline: 16387952]
4. Pauwels R. Global initiative for chronic obstructive lung diseases (GOLD): time to act. Eur Respir J 2001 Dec;18(6):901-902 [FREE Full text] [Medline: 11829093]
5. Whittaker Brown S, Braman S. Recent advances in the management of acute exacerbations of chronic obstructive pulmonary disease. Med Clin North Am 2020 Jul;104(4):615-630. [doi: 10.1016/j.mcna.2020.02.003] [Medline: 32505256]
6. López-Campos JL, Tan W, Soriano JB. Global burden of COPD. Respirology 2016 Jan;21(1):14-23 [FREE Full text] [doi: 10.1111/resp.12660] [Medline: 26494423]
7. Nguyen HT, Collins PF, Pavey TG, Nguyen NV, Pham TD, Gallegos DL. Nutritional status, dietary intake, and health-related quality of life in outpatients with COPD. Int J Chron Obstruct Pulmon Dis 2019;14:215-226 [FREE Full text] [doi: 10.2147/COPD.S181322] [Medline: 30666102]
8. GBD 2015 Chronic Respiratory Disease Collaborators. Global, regional, and national deaths, prevalence, disability-adjusted life years, and years lived with disability for chronic obstructive pulmonary disease and asthma, 1990-2015: a systematic analysis for the Global Burden of Disease Study 2015. Lancet Respir Med 2017 Sep;5(9):691-706 [FREE Full text] [doi: 10.1016/S2213-2600(17)30293-X] [Medline: 28822787]
9. Hartman JE, Boezen HM, de Greef MH, Bossenbroek L, ten Hacken NH. Consequences of physical inactivity in chronic obstructive pulmonary disease. Expert Rev Respir Med 2010 Dec;4(6):735-745. [doi: 10.1586/ers.10.76] [Medline: 21128749]
10. Pitta F, Troosters T, Spruit MA, Probst VS, Decramer M, Gosselink R. Characteristics of physical activities in daily life in chronic obstructive pulmonary disease. Am J Respir Crit Care Med 2005 May 01;171(9):972-977. [doi: 10.1164/rccm.200407-855OC] [Medline: 15665324]
11. Han Y, Heo Y, Hong Y, Kwon SO, Kim WJ. Correlation between physical activity and lung function in dusty areas: results from the chronic obstructive pulmonary disease in dusty areas (CODA) cohort. Tuberc Respir Dis (Seoul) 2019 Oct;82(4):311-318 [FREE Full text] [doi: 10.4046/trd.2019.0006] [Medline: 31172706]
12. Garcia-Aymerich J, Serra I, Gómez FP, Farrero E, Balcells E, Rodríguez DA, PhenotypeCourse of COPD (PAC-COPD) Study Group. Physical activity and clinical and functional status in COPD. Chest 2009 Jul;136(1):62-70. [doi: 10.1378/chest.08-2532] [Medline: 19255291]
13. Jenkins S, Hill K, Cecins NM. State of the art: how to set up a pulmonary rehabilitation program. Respirology 2010 Nov;15(8):1157-1173. [doi: 10.1111/j.1440-1843.2010.01849.x] [Medline: 20920127]
14. Bourbeau J. Making pulmonary rehabilitation a success in COPD. Swiss Med Wkly 2010;140:w13067 [FREE Full text] [doi: 10.4414/smw.2010.13067] [Medline: 20809435]
15. Spruit MA, Pitta F, McAuley E, ZuWallack RL, Nici L. Pulmonary rehabilitation and physical activity in patients with chronic obstructive pulmonary disease. Am J Respir Crit Care Med 2015 Oct 15;192(8):924-933. [doi: 10.1164/rccm.201505-0929CI] [Medline: 26161676]
16. Williams V, Bruton A, Ellis-Hill C, McPherson K. The effect of pulmonary rehabilitation on perceptions of breathlessness and activity in COPD patients: a qualitative study. Prim Care Respir J 2010 Mar;19(1):45-51 [FREE Full text] [doi: 10.4104/pcrj.2009.00044] [Medline: 19603125]
17. Donner C, Muir J. Selection criteria and programmes for pulmonary rehabilitation in COPD patients. Rehabilitation and Chronic Care Scientific Group of the European Respiratory Society. Eur Respir J 1997 Mar;10(3):744-757 [FREE Full text] [Medline: 9073018]
18. Fishman AP. Pulmonary rehabilitation research. Am J Respir Crit Care Med 1994 Mar;149(3 Pt 1):825-833. [doi: 10.1164/ajrccm.149.3.8118655] [Medline: 8118655]
19. Troosters T, Gosselink R, Janssens W, Decramer M. Exercise training and pulmonary rehabilitation: new insights and remaining challenges. Eur Respir Rev 2010 Mar;19(115):24-29 [FREE Full text] [doi: 10.1183/09059180.00007809] [Medline: 20956162]
20. Dürr S, Zogg S, Miedinger D, Steveling EH, Maier S, Leuppi JD. Daily physical activity, functional capacity and quality of life in patients with COPD. COPD 2014 Dec;11(6):689-696. [doi: 10.3109/15412555.2014.898050] [Medline: 24945972]
21. José A, Dal Corso S. Inpatient rehabilitation improves functional capacity, peripheral muscle strength and quality of life in patients with community-acquired pneumonia: a randomised trial. J Physiother 2016 Apr;62(2):96-102 [FREE Full text] [doi: 10.1016/j.jphys.2016.02.014] [Medline: 26996093]
22. Holland AE, Wadell K, Spruit MA. How to adapt the pulmonary rehabilitation programme to patients with chronic respiratory disease other than COPD. Eur Respir Rev 2013 Dec;22(130):577-586 [FREE Full text] [doi: 10.1183/09059180.00005613] [Medline: 24293474]
23. Blumenthal JA, Keefe FJ, Babyak MA, Fenwick CV, Johnson JM, Stott K, et al. Caregiver-assisted coping skills training for patients with COPD: background, design, and methodological issues for the INSPIRE-II study. Clin Trials 2009 Apr;6(2):172-184 [FREE Full text] [doi: 10.1177/1740774509102565] [Medline: 19342470]

XSL•FO
RenderX

24. Maurer J, Rebbapragada V, Borson S, Goldstein R, Kunik ME, Yohannes AM, ACCP Workshop Panel on AnxietyDepression in COPD. Anxiety and depression in COPD: current understanding, unanswered questions, and research needs. Chest 2008 Oct;134(4 Suppl):43S-56S [FREE Full text] [doi: 10.1378/chest.08-0342] [Medline: 18842932]
25. Kessler R, Partridge MR, Miravitlles M, Cazzola M, Vogelmeier C, Leynaud D, et al. Symptom variability in patients with severe COPD: a pan-European cross-sectional study. Eur Respir J 2011 Feb;37(2):264-272 [FREE Full text] [doi: 10.1183/09031936.00051110] [Medline: 21115606]
26. Spruit MA, Singh SJ, Garvey C, ZuWallack R, Nici L, Rochester C, ATS/ERS Task Force on Pulmonary Rehabilitation. An official American Thoracic Society/European Respiratory Society statement: key concepts and advances in pulmonary rehabilitation. Am J Respir Crit Care Med 2013 Oct 15;188(8):e13-e64. [doi: 10.1164/rccm.201309-1634ST] [Medline: 24127811]
27. Bolton CE, Bevan-Smith EF, Blakey JD, Crowe P, Elkin SL, Garrod R, British Thoracic Society Pulmonary Rehabilitation Guideline Development Group, British Thoracic Society Standards of Care Committee. British Thoracic Society guideline on pulmonary rehabilitation in adults. Thorax 2013 Sep;68 Suppl 2:ii1-i30. [doi: 10.1136/thoraxjnl-2013-203808] [Medline: 23880483]
28. Donner C, Ambrosino N, Goldstein RS. Pulmonary Rehabilitation. London: Taylor & Francis group; Aug 04, 2020.
29. Billings PR, Kohn MA, de Cuevas M, Beckwith J, Alper JS, Natowicz MR. Discrimination as a consequence of genetic testing. Am J Hum Genet 1992 Mar;50(3):476-482 [FREE Full text] [Medline: 1539589]
30. Olivier C, Grosbois J, Cortot AB, Peres S, Heron C, Delourme J, et al. Real-life feasibility of home-based pulmonary rehabilitation in chemotherapy-treated patients with thoracic cancers: a pilot study. BMC Cancer 2018 Feb 13;18(1):178 [FREE Full text] [doi: 10.1186/s12885-018-4102-6] [Medline: 29433474]
31. Ryan P, Sawin KJ. The individual and family self-management theory: background and perspectives on context, process, and outcomes. Nurs Outlook 2009;57(4):217-225 [FREE Full text] [doi: 10.1016/j.outlook.2008.10.004] [Medline: 19631064]
32. Sohanpal R. Understanding the reasons for non participation in self-management interventions amongst patients with chronic conditions: addressing and increasing opportunities for patients with advanced chronic obstructive pulmonary disease to access self-management [theses]. Queen Mary University of London. 2015 Apr 29. URL: http://qmro.qmul.ac.uk/xmlui/handle/123456789/15031 [accessed 2015-04-29]
33. Hall C, Nici L, Sood S, ZuWallack R, Castro M. Nonpharmacologic therapy for severe persistent asthma. J Allergy Clin Immunol Pract 2017;5(4):928-935. [doi: 10.1016/j.jaip.2017.04.030] [Medline: 28689843]
34. Ruppel GL, Enright PL. Pulmonary function testing. Respir Care 2012 Jan;57(1):165-175 [FREE Full text] [doi: 10.4187/respcare.01640] [Medline: 22222135]
35. Shah S, Nasb M, Lu M, Huang L, Wang Y, Chen H. Scaling the need, benefits, and risks associated with COVID-19 acute and postacute care rehabilitation: a review. Rehabil Res Pract 2020;2020:3642143 [FREE Full text] [doi: 10.1155/2020/3642143] [Medline: 32908705]
36. Czaja S, Sharit J. Human resource management. In: Designing Training and Instructional Programs for Older Adults. Boca Raton: Taylor & Francis Group; Oct 24, 2012.
37. Moubarak S, Merheb D, Basbous L, Chamseddine N, Bou Zerdan M, Assi HI. COVID-19 and lung cancer: update on the latest screening, diagnosis, management and challenges. J Int Med Res 2022 Sep;50(9):3000605221125047 [FREE Full text] [doi: 10.1177/03000605221125047] [Medline: 36154328]
38. Tabassum T, Rahman A, Araf Y, Ullah MA, Hosen MJ. Management of asthma patients during the COVID-19 pandemic: pathophysiological considerations to address the challenges. Beni Suef Univ J Basic Appl Sci 2022;11(1):20 [FREE Full text] [doi: 10.1186/s43088-022-00204-4] [Medline: 35155689]
39. Gharibo C, Sharma A, Soin A, Shah S, Diwan S, Buenaventura R, et al. Triaging interventional pain procedures during COVID-19 or related elective surgery restrictions: evidence-informed guidance from the American Society of Interventional Pain Physicians (ASIPP). Pain Physician 2020 Aug;23(4S):S183-S204 [FREE Full text] [Medline: 32942785]
40. Coker M, Folayan MO, Michelow IC, Oladokun RE, Torbunde N, Sam-Agudu NA. Things must not fall apart: the ripple effects of the COVID-19 pandemic on children in sub-Saharan Africa. Pediatr Res 2021 Apr;89(5):1078-1086 [FREE Full text] [doi: 10.1038/s41390-020-01174-y] [Medline: 32971527]
41. Bokolo AJ. Exploring the adoption of telemedicine and virtual software for care of outpatients during and after COVID-19 pandemic. Ir J Med Sci 2021 Feb;190(1):1-10 [FREE Full text] [doi: 10.1007/s11845-020-02299-z] [Medline: 32642981]
42. Lukas H, Xu C, Yu Y, Gao W. Emerging telemedicine tools for remote COVID-19 diagnosis, monitoring, and management. ACS Nano 2020 Dec 22;14(12):16180-16193. [doi: 10.1021/acsnano.0c08494] [Medline: 33314910]
43. Tan JY, Conceicao EP, Wee LE, Sim XYJ, Venkatachalam I. COVID-19 public health measures: a reduction in hospital admissions for COPD exacerbations. Thorax 2021 May;76(5):512-513. [doi: 10.1136/thoraxjnl-2020-216083] [Medline: 33273024]
44. Sharma P, Pandey AK, Bhattacharyya DK. Determining crucial genes associated with COVID-19 based on COPD findings. Comput Biol Med 2021 Jan;128:104126 [FREE Full text] [doi: 10.1016/j.compbiomed.2020.104126] [Medline: 33260035]
45. Radzikowska U, Ding M, Tan G, Zhakparov D, Peng Y, Wawrzyniak P, et al. Distribution of ACE2, CD147, CD26, and other SARS-CoV-2 associated molecules in tissues and immune cells in health and in asthma, COPD, obesity, hypertension,

XSL•FO
RenderX

and COVID-19 risk factors. Allergy 2020 Nov;75(11):2829-2845 [FREE Full text] [doi: 10.1111/all.14429] [Medline: 32496587]
46. Sakai T, Hoshino C, Yamaguchi R, Hirao M, Nakahara R, Okawa A. Remote rehabilitation for patients with COVID-19. J Rehabil Med 2020 Sep 04;52(9):jrm00095 [FREE Full text] [doi: 10.2340/16501977-2731] [Medline: 32871014]
47. Wickerson L, Helm D, Gottesman C, Rozenberg D, Singer LG, Keshavjee S, et al. Telerehabilitation for lung transplant candidates and recipients during the COVID-19 pandemic: program evaluation. JMIR Mhealth Uhealth 2021 Jun 17;9(6):e28708 [FREE Full text] [doi: 10.2196/28708] [Medline: 34048354]
48. Brundisini F, Giacomini M, DeJean D, Vanstone M, Winsor S, Smith A. Chronic disease patients' experiences with accessing health care in rural and remote areas: a systematic review and qualitative meta-synthesis. Ont Health Technol Assess Ser 2013;13(15):1-33 [FREE Full text] [Medline: 24228078]
49. Russell TG, Buttrum P, Wootton R, Jull GA. Internet-based outpatient telerehabilitation for patients following total knee arthroplasty. Journal of Bone and Joint Surgery 2011;93(2):113-120 [FREE Full text] [doi: 10.2106/jbjs.i.01375]
50. Roberts P. Staffing an empty schoolhouse: attracting and retaining teachers in rural, remote and isolated communities. ERIC. 2004. URL: https://eric.ed.gov/?id=ED507343 [accessed 2004-09-01]
51. Fitzner K, Moss G. Telehealth--an effective delivery method for diabetes self-management education? Popul Health Manag 2013 Jun;16(3):169-177. [doi: 10.1089/pop.2012.0054] [Medline: 23216062]
52. Martin MA, Catrambone CD, Kee RA, Evans AT, Sharp LK, Lyttle C, Chicago Initiative to Raise Asthma Health Equity Investigative Team. Improving asthma self-efficacy: developing and testing a pilot community-based asthma intervention for African American adults. J Allergy Clin Immunol 2009 Jan;123(1):153-159.e3 [FREE Full text] [doi: 10.1016/j.jaci.2008.10.057] [Medline: 19130936]
53. Brennan D, Mawson S, Brownsell S. Telerehabilitation: enabling the remote delivery of healthcare, rehabilitation, and self management. Stud Health Technol Inform 2009;145:231-248. [Medline: 19592797]
54. Dew A, Bulkeley K, Veitch C, Bundy A, Gallego G, Lincoln M, et al. Addressing the barriers to accessing therapy services in rural and remote areas. Disabil Rehabil 2013 Aug;35(18):1564-1570. [doi: 10.3109/09638288.2012.720346] [Medline: 23009191]
55. Hasson S, Waissengrin B, Shachar E, Hodruj M, Fayngor R, Brezis M, et al. Rapid implementation of telemedicine during the COVID-19 pandemic: perspectives and preferences of patients with cancer. Oncologist 2021 Apr;26(4):e679-e685 [FREE Full text] [doi: 10.1002/onco.13676] [Medline: 33453121]
56. Al Ameen M, Liu J, Kwak K. Security and privacy issues in wireless sensor networks for healthcare applications. J Med Syst 2012 Feb;36(1):93-101 [FREE Full text] [doi: 10.1007/s10916-010-9449-4] [Medline: 20703745]
57. Patel S, Park H, Bonato P, Chan L, Rodgers M. A review of wearable sensors and systems with application in rehabilitation. J Neuroeng Rehabil 2012 Apr 20;9:21 [FREE Full text] [doi: 10.1186/1743-0003-9-21] [Medline: 22520559]
58. Barnett P, Goulding L, Casetta C, Jordan H, Sheridan-Rains L, Steare T, et al. Implementation of telemental health services before COVID-19: rapid umbrella review of systematic reviews. J Med Internet Res 2021 Jul 20;23(7):e26492 [FREE Full text] [doi: 10.2196/26492] [Medline: 34061758]
59. Sampson JP, Kolodinsky RW, Greeno BP. Counseling on the information highway: future possibilities and potential problems. Jour of Counseling & Develop 2011 Dec 23;75(3):203-212 [FREE Full text] [doi: 10.1002/j.1556-6676.1997.tb02334.x]
60. White M, Dorman SM. Receiving social support online: implications for health education. Health Educ Res 2001 Dec;16(6):693-707. [doi: 10.1093/her/16.6.693] [Medline: 11780708]
61. Dolev-Amit T, Leibovich L, Zilcha-Mano S. Repairing alliance ruptures using supportive techniques in telepsychotherapy during the COVID-19 pandemic. Counselling Psychology Quarterly 2020 Jun 14;34(3-4):485-498 [FREE Full text] [doi: 10.1080/09515070.2020.1777089]
62. Simon SR, Evans JS, Benjamin A, Delano D, Bates DW. Patients' attitudes toward electronic health information exchange: qualitative study. J Med Internet Res 2009 Aug 06;11(3):e30 [FREE Full text] [doi: 10.2196/jmir.1164] [Medline: 19674960]
63. Shah SM, Khan RA. Secondary use of electronic health record: opportunities and challenges. IEEE Access 2020;8:136947-136965 [FREE Full text] [doi: 10.1109/access.2020.3011099]
64. Staring K. Organizational open source in the Global South [PhD thesis]. University of Oslo. 2011 May. URL: http://tinyurl.com/4bbfvt9u [accessed 2024-02-14]
65. Kichloo A, Albosta M, Dettloff K, Wani F, El-Amir Z, Singh J, et al. Telemedicine, the current COVID-19 pandemic and the future: a narrative review and perspectives moving forward in the USA. Fam Med Community Health 2020 Aug;8(3):e000530 [FREE Full text] [doi: 10.1136/fmch-2020-000530] [Medline: 32816942]
66. Tyree PT, Lind BK, Lafferty WE. Challenges of using medical insurance claims data for utilization analysis. Am J Med Qual 2006;21(4):269-275 [FREE Full text] [doi: 10.1177/1062860606288774] [Medline: 16849784]
67. Ho Chan WS. Taiwan's healthcare report 2010. EPMA J 2010 Dec;1(4):563-585 [FREE Full text] [doi: 10.1007/s13167-010-0056-8] [Medline: 23199110]
68. Health financing challenges and institutional options to move towards universal coverage in Nicaragua. World Health Report. 2010. URL: https://cdn.who.int/media/docs/default-source/health-financing/technical-briefs-background-papers/24nicaragua.pdf?sfvrsn=1cd08511_3&download=true [accessed 2021-06-11]

XSL•FO
RenderX

69. Lahham A, McDonald CF, Mahal A, Lee AL, Hill CJ, Burge AT, et al. Home-based pulmonary rehabilitation for people with COPD: A qualitative study reporting the patient perspective. Chron Respir Dis 2018 May;15(2):123-130 [FREE Full text] [doi: 10.1177/1479972317729050] [Medline: 28868892]
70. Ensor T, Cooper S. Overcoming barriers to health service access: influencing the demand side. Health Policy Plan 2004 Mar;19(2):69-79. [doi: 10.1093/heapol/czh009] [Medline: 14982885]
71. Bhatt SP, Patel SB, Anderson EM, Baugh D, Givens T, Schumann C, et al. Video telehealth pulmonary rehabilitation intervention in chronic obstructive pulmonary disease reduces 30-day readmissions. Am J Respir Crit Care Med 2019 Aug 15;200(4):511-513 [FREE Full text] [doi: 10.1164/rccm.201902-0314LE] [Medline: 30978302]
72. Bhatt SP, Baugh D, Hitchcock J, Kim Y, Cutter G, Aban I, et al. Video telehealth pulmonary rehabilitation for chronic obstructive pulmonary disease is associated with clinical improvement similar to center-based pulmonary rehabilitation. Ann Am Thorac Soc 2022 Feb;19(2):331-333 [FREE Full text] [doi: 10.1513/AnnalsATS.202104-419RL] [Medline: 34388079]
73. Brownstein JS, Freifeld CC, Madoff LC. Digital disease detection--harnessing the web for public health surveillance. N Engl J Med 2009 May 21;360(21):2153-5, 2157 [FREE Full text] [doi: 10.1056/NEJMp0900702] [Medline: 19423867]
74. Lavin B, Dormond C, Scantlebury MH, Frouin P, Brodie MJ. Bridging the healthcare gap: Building the case for epilepsy virtual clinics in the current healthcare environment. Epilepsy Behav 2020 Oct;111:107262 [FREE Full text] [doi: 10.1016/j.yebeh.2020.107262] [Medline: 32645620]
75. Delitto A, Erhard RE, Bowling RW. A treatment-based classification approach to low back syndrome: identifying and staging patients for conservative treatment. Phys Ther 1995 Jun;75(6):470-485. [doi: 10.1093/ptj/75.6.470] [Medline: 7770494]

Abbreviations

COPD: chronic obstructive pulmonary disease
PR: pulmonary rehabilitation

Edited by A Kushniruk; submitted 22.07.23; peer-reviewed by HH Chen, UK Chalwadi; comments to author 20.11.23; revised version received 01.12.23; accepted 24.01.24; published 07.03.24.

Please cite as:
Chien SY, Wong AMK, Tseng W, Hu HC, Cho HY
Feasibility and Design Factors for Home-Based Pulmonary Rehabilitation of Patients With Chronic Obstructive Pulmonary Disease and Chronic Lung Diseases Based on a People-Object-Environment Framework: Qualitative Interview Study
JMIR Hum Factors 2024;11:e51150
URL: https://humanfactors.jmir.org/2024/1/e51150
doi:10.2196/51150
PMID:38452366

©Shih-Ying Chien, Alice May-Kuen Wong, Winston Tseng, Han-Chung Hu, Hsiu-Ying Cho. Originally published in JMIR Human Factors (https://humanfactors.jmir.org), 07.03.2024. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on https://humanfactors.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

A Technological Tool Aimed at Self-Care in Patients With Multimorbidity: Cross-Sectional Usability Study

Rodrigo Medina-García[1,2,3,4*], MD; Juan A López-Rodríguez[1,2,4,5,6*], MSc, MD, PhD; Cristina María Lozano-Hernández[1,3,4,5*], MSc; Verónica Ruiz Bejerano[7*], MSc; Paride Criscio[8*]; Isabel Del Cura-González[1,5,6*], MD, PhD; TeNDER Atención Primaria Group[1*]

[1]Primary Care Research Unit, Madrid Health Service, Madrid, Spain
[2]General Ricardos Primary Health Care Centre, Madrid Health Service, Madrid, Spain
[3]Interuniversity Doctoral Program in Epidemiology and Public Health, Rey Juan Carlos University, Alcorcón, Madrid, Spain
[4]Biosanitary Research and Innovation Foundation of Primary Care, Madrid, Spain
[5]Research Network on Chronicity, Primary Care and Health Promotion, Madrid, Spain
[6]Department of Medical Specialties and Public Health, Faculty of Health Sciences, Rey Juan Carlos University, Alcorcón, Madrid, Spain
[7]Visual Telecommunications Application Research Group, Universidad Politécnica de Madrid, Madrid, Spain
[8]DataWizard Srl, Rome, Italy
[*]all authors contributed equally

Corresponding Author:
Rodrigo Medina-García, MD
General Ricardos Primary Health Care Centre
Madrid Health Service
Calle General Ricardos, 131
Madrid, 28019
Spain
Phone: 34 660802663
Email: rodrigo.medina@salud.madrid.org

Abstract

Background: Information and communication technologies (ICTs) have been positioned as useful tools to facilitate self-care. The interaction between a patient and technology, known as usability, is particularly important for achieving positive health outcomes. Specific characteristics of patients with chronic diseases, including multimorbidity, can affect their interaction with different technologies. Thus, studying the usability of ICTs in the field of multimorbidity has become a key element to ensure their relevant role in promoting self-care.

Objective: The aim of this study was to analyze the usability of a technological tool dedicated to health and self-care in patients with multimorbidity in primary care.

Methods: A descriptive observational cross-sectional usability study was performed framed in the clinical trial in the primary care health centers of Madrid Health Service of the TeNDER (Affective Based Integrated Care for Better Quality of Life) project. The TeNDER technological tool integrates sensors for monitoring physical and sleep activity along with a mobile app for consulting the data collected and working with self-management tools. This project included patients over 60 years of age who had one or more chronic diseases, at least one of which was mild-moderate cognitive impairment, Parkinson disease, or cardiovascular disease. From the 250 patients included in the project, 38 agreed to participate in the usability study. The usability variables investigated were effectiveness, which was determined by the degree of completion and the total number of errors per task; efficiency, evaluated as the average time to perform each task; and satisfaction, quantified by the System Usability Scale. Five tasks were evaluated based on real case scenarios. Usability variables were analyzed according to the sociodemographic and clinical characteristics of patients. A logistic regression model was constructed to estimate the factors associated with the type of support provided for task completion.

Results: The median age of the 38 participants was 75 (IQR 72.0-79.0) years. There was a slight majority of women (20/38, 52.6%) and the participants had a median of 8 (IQR 7.0-11.0) chronic diseases. Thirty patients completed the usability study, with a usability effectiveness result of 89.3% (134/150 tasks completed). Among the 30 patients, 66.7% (n=20) completed all tasks and 56.7% (17/30) required personalized help on at least one task. In the multivariate analysis, educational level emerged

XSL•FO
RenderX

as a facilitating factor for independent task completion (odds ratio 1.79, 95% CI 0.47-6.83). The median time to complete the total tasks was 296 seconds (IQR 210.0-397.0) and the median satisfaction score was 55 (IQR 45.0-62.5) out of 100.

Conclusions: Although usability effectiveness was high, the poor efficiency and usability satisfaction scores suggest that there are other factors that may interfere with the results. Multimorbidity was not confirmed to be a key factor affecting the usability of the technological tool.

Trial Registration: Clinicaltrials.gov NCT05681065; https://clinicaltrials.gov/study/NCT05681065

(JMIR Hum Factors 2024;11:e46811) doi:10.2196/46811

KEYWORDS

user-centered design; multimorbidity; comorbid; self-care; medical informatics; primary health care; chronic disease; chronic condition; chronic illness; primary care; usability; telemedicne; telehealth; information and communication technologies; ICT; digital health; eHealth; human-computer interaction

Introduction

Multimorbidity, which is generally defined as the presence of two or more simultaneous chronic diseases in a patient, is a major challenge for health systems [1]. In the European Union, up to 50 million people are estimated to have multimorbidity [2]. Barnett et al [3] estimated a multimorbidity prevalence of 64.9% among patients aged 65-84 years and of 81.5% for those 85 years or older [3]. In recent years, patient-centered care models [4] and, more specifically, interventions aimed at self-care education have made it possible to optimize how patients with multimorbidity manage their chronic diseases [5]. This type of intervention makes it easier for patients to identify their health-problem needs and to identify techniques that can help them make decisions, take appropriate actions, and modify them as they present changes in their diseases [6].

Information and communication technologies (ICTs) have been positioned as useful tools to facilitate self-care [7]. The different self-care strategies using ICTs include those dedicated to the monitoring of biometric parameters through wearable technologies or portable devices and mobile apps [8,9]. To achieve positive health outcomes from these interventions, the interaction between a patient and technology is particularly important. The description of this interaction between the technology, the specific tasks to be developed, and the end user is a property known as usability [10].

Research on usability has grown in parallel with the development of ICTs in health [11]. Reports from international organizations such as the 2012-2020 eHealth Action Plan of the European Commission [12] and the World Health Organization Global Strategy on Digital Health 2020-2025 [13] summarized the importance of the development of technological tools that take into account their interaction with the special conditions of older adults. The study of usability can help determine the reasons for low patient adherence and adoption of a specific technological tool. The improvements in usability could facilitate interaction through several mechanisms: reducing anxiety related to the use of new tools, increasing accessibility and distribution among a greater number of users, and reducing the possible risks derived from misuse [10].

In evaluating usability, the International Organization for Standardization (ISO) 9126 standard [14] assesses the quality of the product [15] and the ISO 9241 standard focuses on processes, referred to as "the extent to which users in a specific environment can use a product to achieve objectives of effectiveness, efficiency and satisfaction in a particular task" [16]. Thus, usability comprises the effectiveness in usability, defined as the degree of completion [11,17,18] and the total number of errors per task; efficiency in usability, defined as the average time to perform the task; and satisfaction in usability, defined as the degree to which the user's physical and emotional responses resulting from the use of a product satisfy their needs and expectations. The most commonly used methods for usability evaluation are questionnaires and interviews carried out after the use of the technological tool for a certain period of time.

The systematic reviews of Saeed et al [17] and Zapata et al [18] analyzed the definitions of the ISO standards and the methods used in evaluating the usability of health-related technological tools. Their results indicate that the most frequent usability problems are those related to visual aspects of the system and the ability to learn and use specific features [17,18]. However, because the results are limited to a specific technology and may not be generalizable, their interpretation should take into account the special characteristics of the end users, including their health conditions.

The specific characteristics of patients with chronic diseases can affect their interaction with different technologies. For example, in the usability evaluation studies of an app for diabetes self-management [19,20] and that conducted on an automatic drug dispenser for patients with dementia [21], characteristics of the patients were identified that interact with different aspects of usability. Relatedly, Wildenbos et al [22] differentiated four traits related to aging and chronic diseases that act as barriers to usability: cognitive, physical, motivational, and perceptual. Although the development of technological tools aimed at self-care is increasing, their usability has thus far mainly been evaluated in patients with specific isolated pathologies such as in the previous examples. Research from the perspective of patients with multimorbidity has been increasing in recent years [23,24] but remains insufficient [25], even though multimorbidity is the most common way of reporting chronic diseases in the population over 60 years of age [3].

Thus, studying the usability of ICTs has become a key element in the field of multimorbidity [26] to ensure its relevant role in

XSL•FO
RenderX

promoting self-care [7]. Along these lines, the TeNDER (Affective Based Integrated Care for Better Quality of Life) project [27,28] was a multisectoral project funded by Horizon 2020, the EU Framework Programme for research and innovation. The TeNDER project developed an integrated care model to manage multimorbidity in patients with dementia, Parkinson disease, and cardiovascular disease in four European countries: Spain, Germany, Italy, and Slovenia. One of the clinical studies related to the TeNDER project was a multicenter, randomized, parallel-group clinical trial carried out in Spain with the main objective of evaluating the effectiveness of the TeNDER system to improve quality of life in patients with chronic diseases. Secondary aims were to describe the satisfaction of patients and their caregivers and the usability of the TeNDER system [29].

The objective of this study was to analyze the usability of a technological tool (TeNDER) dedicated to health and self-care in patients with multimorbidity in primary care.

Methods

Design

This was a descriptive observational cross-sectional study of usability. This study was framed in the clinical trial in the primary care health centers of Madrid Health Service of the TeNDER project (ClinicalTrials.gov NCT05681065) [29].

Ethical Considerations

This study respects the basic ethical principles of autonomy, beneficence, justice, and nonmaleficence, and its development followed the norms of Good Clinical Practice and the principles enunciated in the latest Declaration of Helsinki (Seoul 2013). The study obtained a favorable report from the Research Ethics Board of the Hospital Universitario 12 de Octubre (20/450) and was approved by the Central Research Commission of the Community of Madrid (PC:39/20). Informed consent was obtained from all participants involved in the study. No camera recording or any other identification was made. Patients were included with an anonymous identifier in the data collection logbook (DCL). All data were processed based on the provisions of the EU General Data Protection Regulation 2016/679 of the European Parliament and the Council (April 27, 2016) and the Organic Law on Data Protection and Guarantee of Digital Rights in the Spanish territory (LOPDGDD 3/2018 of 5 December). Participants did not receive any financial compensation for their participation in the study. The only compensation was that received through the user experience during the use of the technological tool.

Population and Sample

The study population included patients with one or more chronic diseases recruited from four primary care health centers in the Community of Madrid that had been included in the TeNDER project by their referring professionals [27,28].

Patients over 60 years of age who had visited their health center in the last year and who had any of the following chronic diseases were included: mild-moderate cognitive impairment (Mini-Mental State Evaluation [MMSE] score 19-28 points), Parkinson disease, or cardiovascular disease, which includes patients with heart failure, chronic ischemic heart disease, or atrial fibrillation. Patients with a life expectancy of less than 6 months based on the opinion of their health care professionals, severe mental illness, incapacity for autonomous movement, or an MMSE score of less than 19 points were excluded.

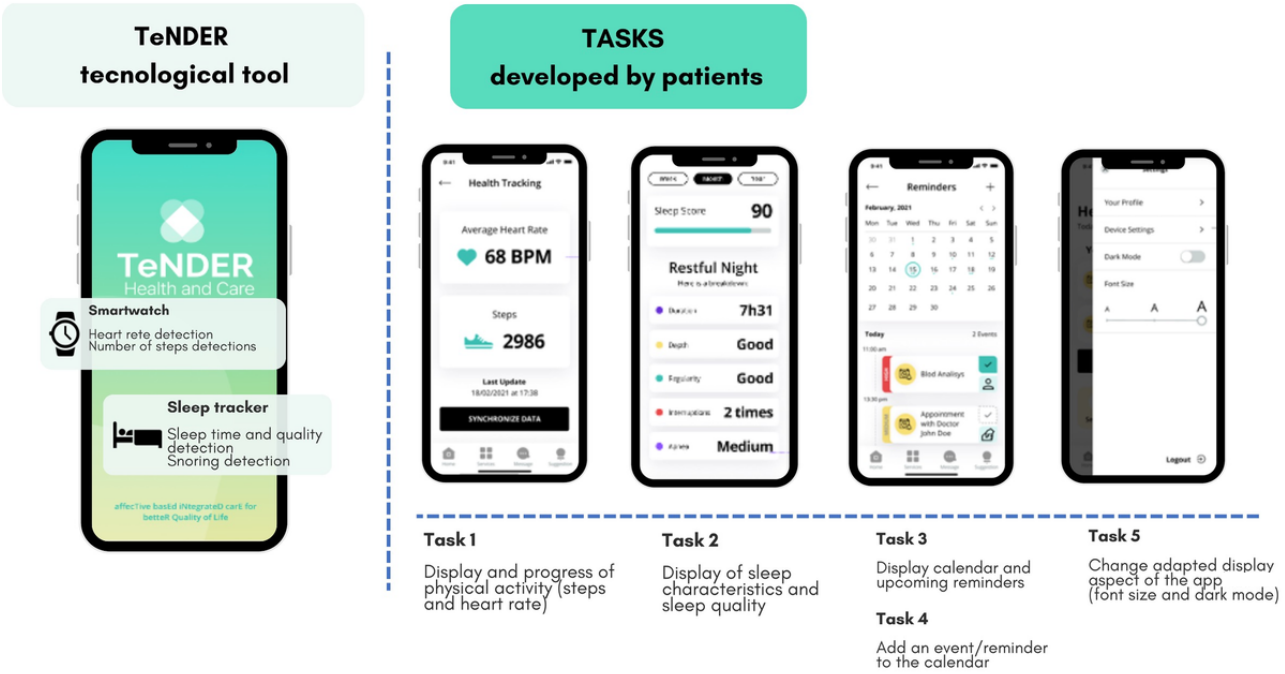
The 250 patients included in the primary care health centers of Madrid Health Service for the TeNDER project were invited to participate via text messages with a mobile instant messaging app. Thirty-eight patients with multimorbidity (≥2 chronic diseases) agreed to participate. Considering an approximate 90% completion rate of tasks in previous usability evaluation studies of monitoring tools [19,25,30], with this sample size, we report a precision of 9.6% with the 95% CI.

TeNDER Technological Tool

The TeNDER technological tool is a web-based platform that included integrating sensors such as a smartwatch for monitoring physical activity, a sleep tracker to study sleep activity, and a mobile app in which the data collected are displayed and tools for self-management are offered (Figure 1). All of the TeNDER ecosystem technology was developed through a co-design process with all relevant stakeholders using a patient-centered approach. During the project, the functionalities and the mobile app were validated and released after user validation within an incremental development approach, ensuring a feedback framework that provided iterative refinement and improvements of the mobile app.

XSL•FO
RenderX

Figure 1. TeNDER technological tool and tasks to be performed by patients for the usability evaluation.



Variables

The main variable of the study was usability effectiveness, which was determined by the degree of completion and the total number of errors per task. Five tasks were designed using the TeNDER system (Textbox 1). The tasks to be evaluated were based on real case scenarios to simulate how patients would interact with the system in a real-life situation according to the care and self-management process based on the main functionalities of the TeNDER app [29]. The tasks were validated by a panel of three health professionals with experience in the study of usability to verify the accuracy of the content and the context. The degree of completion of the tasks was coded using three categories: (0) not completed when the subject was unable to complete the task (inability to progress or to request advanced help or interruption in task execution), (1) completed with personalized help in the form of comments or directed indications, and (2) completed independently when the user was able to carry out the task either without any help from the person in charge of the test or with the aid of minor indications. An error was coded when the subject made errors that could not be solved or that prevented further progress.

Textbox 1. Tasks evaluated for the usability study.

Task 1: Display and progress of physical activity (steps and heart rate).
Task 2: Display of sleep characteristics and sleep quality.
Task 3: Display calendar and upcoming reminders.
Task 4: Add an event/reminder to the calendar.
Task 5: Change-adapted display aspects of the app (font size and dark mode).

As secondary variables, usability efficiency was determined by timing each individual task and calculating the average time in each task. Usability satisfaction was quantified by administering the System Usability Scale (SUS) (Multimedia Appendix 1) in its Spanish-validated version [31]. For this scale, the global score ranges from 0 to 100, where higher values indicate greater usability satisfaction. According to Bangor et al [32], SUS scores of 70-100 indicate acceptable, whereas scores of 0-50 indicate not acceptable; scores between 50 and 70 are considered to indicate marginally acceptable results.

Sociodemographic variables collected included age, sex, and education level, and clinical variables included type and number of chronic diseases. Chronic pathologies were identified according to the proposals in the O'Halloran classification [33] (Multimedia Appendix 2).

Technology-related variables included previous use of touch screens and the affinity for technology interaction (ATI) scale [34]. For this scale, the global score ranges between 1 and 6, where higher values indicate a greater affinity for the technology (Multimedia Appendix 3).

Data Collection

The variables were collected by interview with the patient in consultation with their referring professional and were recorded in an electronic DCL designed ad hoc with the Research Electronic Data Capture (REDCap) tool hosted on the secure storage server of the institution. REDCap is a secure, web-based

XSL•FO
RenderX

software platform designed to support data capture for research studies [35,36].

The patients received the TeNDER technological tool. The usability study was carried out 48 hours afterward based on the execution of tasks in a face-to-face session with a member of the research team who could provide assistance. To record the variables that measure usability, a real-time screen recording of the mobile device was performed during the entirety of task performance. One member of the research team analyzed the recordings. The start and end times were determined from the time the instructions were offered until the moment each task was completed; that information was subsequently transferred to the DCL.

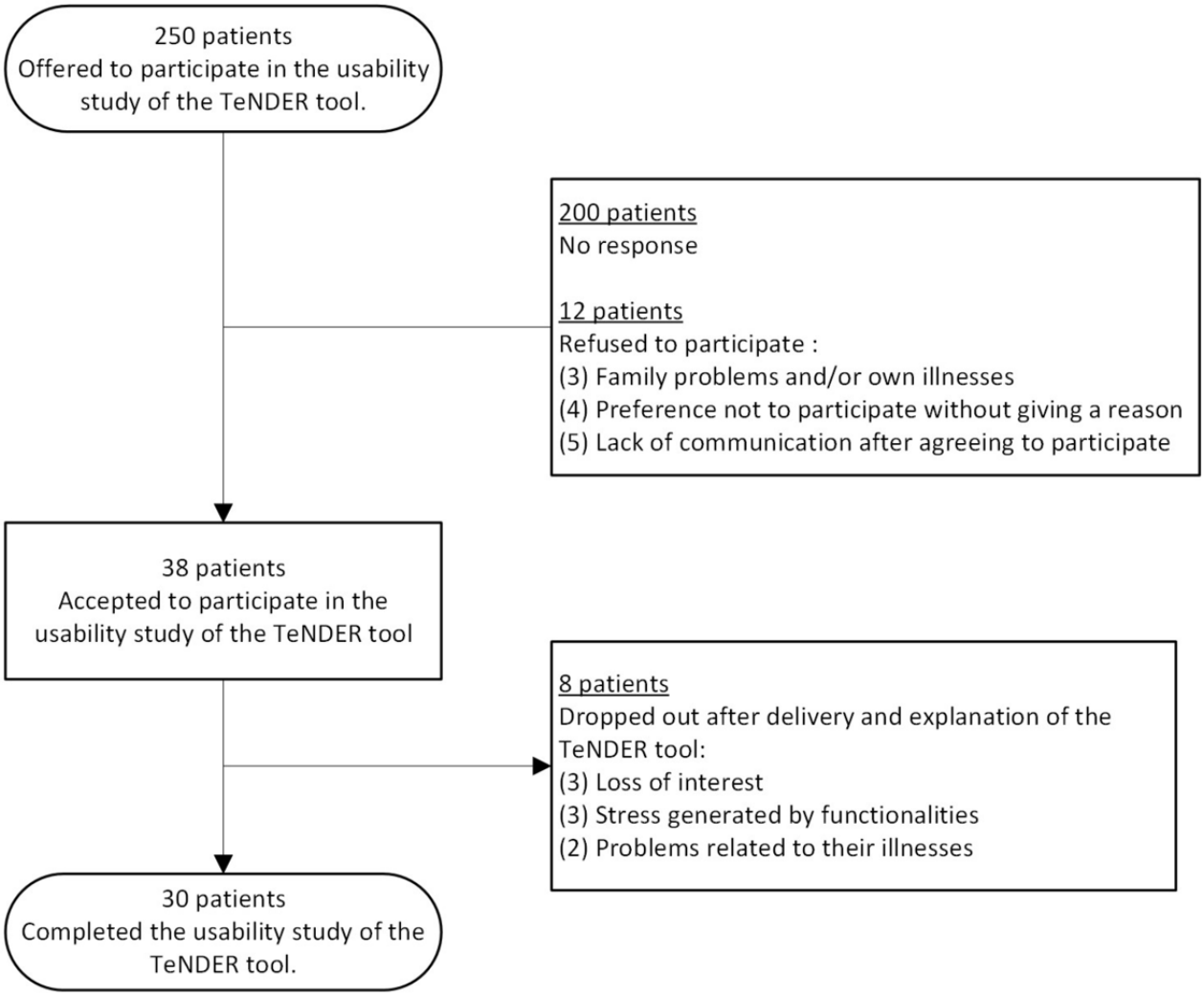
Statistical Analysis

The categorical variables are described as frequencies and percentages. The quantitative variables are described as medians and IQR, as they were nonnormally distributed for the number of patients under study. The main result variable was the proportion of completed tasks (usability effectiveness) with its 95% CI. As secondary outcome variables, the mean effective time to perform each of the tasks (usability efficiency) and the mean score in the SUS questionnaire (usability satisfaction) were estimated. The association of the different usability components (efficacy, effectiveness, and satisfaction) with the sociodemographic and clinical variables was evaluated using the χ^2 test for categorical variables and the Student *t* test for quantitative variables (the Mann-Whitney *U* test was used for comparison of variables that did not follow a normal distribution). The factors associated with completing the task in an independent manner were analyzed using a multiple logistic regression model with robust estimators. The dependent variable was completing the task autonomously. The independent variables were those found to be statistically significant in the bivariable analyses or variables that are otherwise considered to be clinically important. STATA 14 software was used for all statistical analyses.

Results

Among the 250 patients included in the TeNDER project invited, 38 (15.2%) agreed to participate in this study (Figure 2). There were no differences in sociodemographic characteristics between those who refused to participate and the final sample. Finally, 30 patients completed the usability study.

Figure 2. Flowchart of the participants.



XSL•FO
RenderX

The median age of the participants was 75 (IQR 72.0-79.0) years and 20/38 (52.6%) were women. With a median of 8 (IQR 7.0-11.0) chronic diseases, 89.5% of the patients had at least one cardiovascular risk factor. Syndromes that include anxiety and depression occurred at significantly different rates between women (11/20, 55%) and men (3/18, 16.7%). A total of 83.8% of the patients had previously interacted with touch screens, and the median result on the ATI scale was 3.4 (IQR 3.0-3.8), which differed between men and women.

The baseline characteristics of the patients are presented in Table 1.

Thirty patients completed the usability study. Among them, 66.7% (20/30) completed all the tasks to be evaluated. All patients were able to complete at least three of the five proposed tasks and 10 patients did not complete at least one task. At least one mistake was made while carrying out the tasks in 28/30 patients. A total of 66.7% (10/15) of the women required personalized help in at least one of the tasks for their completion. The median usability satisfaction in the SUS questionnaire was 55 (IQR 45.0-62.5).

A total of 150 tasks were carried out among all users and 89.3% (134/150) of the tasks were completed. Tasks 1 and 2 were completed 100% (60/60) of the time. Task 4 was completed at a lower proportion than the other tasks (22/30, 73.3%) and presented the highest number of errors (mean 2.5, SD 0.47). Task 3 and task 4 required personalized help to be completed (10/30, 33.3% and 8/30, 26.7%, respectively). The results of the different usability components are shown in Table 2 (also see Multimedia Appendix 4) and the details according to the different tasks are shown in Table 3 (also see Multimedia Appendix 5). The results of the subgroup usability analysis considering patients with cognitive deficits are provided in Multimedia Appendix 6.

In the multivariate analysis of the characteristics of the total tasks evaluated that were completed (Table 4), education level emerged as a facilitating factor to complete the task in an independent manner. Being male, having diseases related to cognition, and age hindered the completion of the task without help, with the latter factor being statistically significant.

XSL•FO
RenderX

Table 1. Patient baseline characteristics.

Characteristics	Total	Women	Men	*P* value
Participants, n (%)	38 (100)	20 (52.6)	18 (47.4)	N/A[a]
Age (years), median (IQR)	75.0 (72.0-79.0)	75.0 (69.0-80.0)	76.0 (73.0-78.0)	.67
Education level, n (%)				.19
Up to primary studies	17 (44.7)	8 (40.0)	9 (50.0)	
Secondary studies	7 (18.4)	6 (30.0)	1 (5.6)	
Higher education	14 (36.8)	6 (30.0)	8 (44.4)	
Number of chronic diseases, median (IQR)	8.0 (7.0-11.0)	9.0 (8.0-11.5)	8.0 (5.0-10.0)	.05
Cardiovascular risk factors, n (%)				
Total	34 (89.5)	19 (95.0)	15 (83.3)	.33
Arterial hypertension	24 (63.2)	13 (65.0)	11 (61.1)	>.99
Lipid metabolism disorders	24 (63.2)	16 (80.0)	8 (44.4)	.04
Type 2 diabetes mellitus	13 (34.2)	8 (40.0)	5 (27.8)	.51
Overweight/obesity	13 (34.2)	7 (35.0)	6 (33.3)	>.99
Cardiovascular disease	28 (73.7)	4 (20.0)	4 (22.2)	>.99
Perception problems, n (%)				
Total	21 (55.3)	9 (45.0)	12 (66.7)	.21
Vision problems	16 (42.1)	8 (40.0)	8 (44.4)	>.99
Musculoskeletal problems, n (%)	32 (84.2)	19 (95.0)	13 (72.2)	.08
Cognition problems, n (%)				
Total	21 (55.3)	13 (65.0)	8 (44.4)	.33
Cognitive impairment	9 (23.7)	5 (25.0)	4 (22.2)	>.99
Anxiety-depression	14 (36.8)	11 (55.0)	3 (16.7)	.02
Sleep disorders, n (%)	12 (31.6)	4 (20.0)	8 (44.4)	.16
Previous interaction with touch screens, n (%)	31 (83.8)	15 (78.9)	16 (88.9)	.66
Affinity for technology interaction scale, median (IQR)	3.4 (3.0-3.8)	3.3 (2.7-3.8)	3.6 (3.2-3.9)	.16

[a]N/A: not applicable.

XSL•FO
RenderX

Table 2. Usability results according to the total number of patients who completed the study and total number of tasks completed.

Usability metric	Total	≤10 CDs[a]	>10 CDs	Women	Men
Per patient					
Number of patients	30	16	14	15	15
Usability effectiveness					
Number of tasks completed, median (IQR)	5 (4.0-5.0)	5 (4.0-5.0)	5 (4.0-5.0)	5 (4.0-5.0)	5 (5.0-5.0)
At least one task with personalized help, n (%)	17 (56.7)	9 (56.2)	8 (57.1)	10 (66.7)	7 (46.7)
All tasks completed independently, n (%)	12 (40.0)	6 (37.5)	6 (42.8)	5 (33.3)	7 (46.7)
Number of errors made, median (IQR)	4.0 (2.0-8.0)	5 (2.0-9.0)	4 (2.0-6.0)	6 (3.0-9.0)	3.5 (1.0-5.0)
Usability efficiency: time to complete all tasks (seconds), median (IQR)	296.0 (210.0-397.0)	300.0 (236.0-397.0)	284.0 (205.0-431.0)	296.0 (201.0-350.0)	293.5 (211.0-447.0)
Usability satisfaction: SUS[b] questionnaire score, median (IQR)	55.0 (45.0-62.5)	50.0 (45.0-58.8)	61.2 (42.5-70.0)	52.5 (40.0-62.5)	60.0 (47.5-67.5)
Tasks					
Number of tasks	150	80	70	75	75
Usability effectiveness					
Proportion of tasks completed, n (%)	134 (89.3)	71 (88.7)	63 (90.0)	65 (86.7)	69 (92.0)
Number of errors per task, mean (SD)	1.0 (1.7)	1.2 (2.0)	0.8 (1.3)	1.4 (2.0)[c]	0.6 (1.2)[c]
Usability efficiency: time per task (seconds), mean (SD)	65.4 (92.7)	66.8 (95.1)	63.8 (90.7)	60.1 (79.3)	70.9 (105.1)

[a]CD: chronic disease.

[b]SUS: System Usability Scale.

[c]P=.007. This is the only comparison in which significant differences were found. The table with all P values is provided in Multimedia Appendix 4.

XSL•FO
RenderX

Table 3. Usability results by task.[a]

Usability metric	Total (N=30)	Women (n=15)	Men (n=15)	≤10 CDs[b] (n=16)	>10 CDs (n=14)	Up to secondary education (n=18)	Postsecondary education (n=12)
Task 3							
Usability effectiveness							
Number of patients completing the task, n (%)	28 (93.3)	14 (93.3)	14 (93.3)	14 (87.5)	14 (100.0)	16 (88.9)	12 (100.0)
Completed the task with personalized help, n (%)	10 (33.3)	5 (33.3)	5 (33.3)	5 (31.2)	5 (35.7)	7 (38.9)	3 (25.0)
Number of errors made, median (IQR)	0.0 (0.0-1.0)	0.0 (0.0-1.0)	0.0 (0.0-1.0)	1.0 (0.0-1.5)	0.0 (0.0-1.0)	0.0 (0.0-1.0)	0.5 (0.0-2.0)
Usability efficiency: time to perform the task (seconds), median (IQR)	57.0 (33.0-90.0)	40.0 (30.0-70.0)	65.0 (33.0-105.0)	50.0 (23.0-90.0)	58.5 (37.0-90.0)	60.0 (30.0-90.0)	48.5 (34.0-80.0)
Task 4							
Usability effectiveness							
Number of patients completing the task, n (%)	22 (73.3)	9 (60.0)	13 (86.7)	12 (75.0)	10 (71.4)	10 (55.6)[d]	12 (100.0)[d]
Completed the task with personalized help, n (%)	8 (26.7)	4 (26.7)	4 (26.7)	4 (25.0)	4 (28.6)	5 (27.8)	3 (25.0)
Number of errors made, median (IQR)	2.0 (0.0-4.0)	3.0 (1.0-5.0)[e]	1.0 (0.0-3.0)[e]	2.0 (0.0-4.0)	1.5 (0.0-3.0)	3.0 (0.0-5.0)	1.0 (0.0-2.0)
Usability efficiency: time to perform the task (seconds), median (IQR)	182.5 (0.0-280.0)	185.0 (0.0-220.0)	180.0 (137.0-300.0)	186.0 (80.0-265.0)	174.0 (0.0-280.0)	176.0 (0.0-300.0)	186.0 (161.5-225.0)
Task 5							
Usability effectiveness							
Number of patients completing the task, n (%)	26 (86.7)	12 (80.0)	14 (93.3)	15 (93.8)	11 (78.6)	14 (77.8)	12 (100.0)
Completed the task with personalized help, n (%)	5 (16.7)	2 (13.3)	3 (20.0)	3 (18.8)	2 (14.3)	4 (22.2)	1 (8.3)
Number of errors made, median (IQR)	1.0 (0.0-2.0)	1.0 (1.0-3.0)[c]	0.0 (0.0-1.0)[c]	0.5 (0.0-2.0)	1.0 (0.0-2.0)	1.5 (0.0-3.0)	0.5 (0.0-1.0)
Usability efficiency: time to perform the task (seconds), median (IQR)	20.0 (10.0-50.0)	30.0 (6.0-75.0)	18.0 (10.0-21.0)	20.0 (14.0-75.0)	19.0 (6.0-31.0)	20.5 (10.0-80.0)	19.0 (12.0-27.5)

[a]The usability results for tasks 1 and 2 are not included because they showed 100% effectiveness in usability; *P* values are only indicated for comparisons in which significant differences were found. The table with all *P* values is provided in Multimedia Appendix 5.

[b]CD: chronic disease.

[c]*P*=.01.

[d]*P*=.01.

[e]*P*=.04.

XSL•FO
RenderX

Table 4. Factors associated with completing a task in an independent manner.

Associated factors	Odds ratio (95% CI)	*P* value
Male sex	0.81 (0.24-2.74)	.74
Age	0.85 (0.77-0.94)	.002
Education level		
Up to secondary education	Reference	N/A[a]
Postsecondary education	1.79 (0.47-6.83)	.39
Diseases related to cognition	0.18 (0.04-0.81)	.03

[a]N/A: not applicable.

Discussion

Main Findings

The usability effectiveness of the TeNDER technological tool was 89.3% (134/150). Overall, 40% (12/30) of the patients completed all tasks independently. Task 4 was completed at a lower proportion than the rest of the tasks (22/30, 73.3%) and presented the highest number of errors (mean 2.5, SD 0.47). The usability efficiency, evaluated as the median time to complete the total tasks, was 296.0 seconds (IQR 210.0-397.0), with an average value per task of 65.4 seconds (SD 92.7). The satisfaction in usability perceived by the patients was acceptable (mean 52.2, SD 16.9). Being male, having diseases related to cognition, and age were factors that hindered the completion of the task without personalized help, among which only age was statistically significant.

Comparison With Other Studies

The usability effectiveness of the TeNDER technological tool was 89.3%, which is similar to the results of previous studies carried out on different categories of patients for similar technological developments. Sánchez-Morillo et al [30] evaluated the usability of a technological tool aimed at monitoring the symptoms of patients with chronic obstructive pulmonary disease, and Georgsson et al [19] evaluated a system designed for the management of self-care in patients with diabetes. The proportion of tasks completed was 88% and 91%, respectively, despite the opposite characteristics of the participants with respect to the level of education and affinity for technology in each of the studies. As in our study, the degree of task complexity could have been adapted to the characteristics of the potential users: tasks 1 and 2 were completed by all patients, whereas the rest of the tasks, of greater complexity, were completed by only those with higher education. For those who did not have higher education, the task completion rate reached up to 55.6%. These differences in the use of technology depending on the level of education have been confirmed in previous studies [37].

Despite the high proportion of completed tasks, 56.7% of the patients required personalized help to complete at least one of the tasks. Older age and cognition-related diseases were risk factors for requiring personalized help to complete the tasks. Previous experience in evaluating the usability of a computerized system for self-care management aimed at patients with chronic diseases yielded similar percentages of effectiveness in usability and help for task completion [25], which points to the importance of having family members or professionals assist patients with chronic diseases to interact with a mobile app [38].

The median value for usability satisfaction was 55.0 (IQR 45.0-62.5), which is a low marginal score over not acceptable [32]. Ligons et al [21] obtained similar results and indicated that the degree of response in satisfaction with a technology or system may not be related to the ability for the completion of its tasks. That is, patients may be able to complete tasks without knowing why they have completed them or how they can benefit from them in their day-to-day lives. Other studies, including that of Sánchez-Morillo et al [30], suggested that high levels of satisfaction may be caused by the presence of qualified professionals who assisted during the usability evaluation.

The median age of the patients in our study was 75 years and a high degree of multimorbidity was notable, with a median of up to eight chronic diseases. Previous studies have analyzed the usability of a technology from the perspective of patients with a chronic index disease in particular [19,20,30,39]. For example, Wildenbos et al [22] analyzed how chronic diseases can affect the usability of technological tools. Thus, a single chronic disease can be the cause of physical, cognitive, and perception barriers [22]. Medical conditions that could favor the appearance of these barriers are represented in our study: diabetes, cardiovascular disease, cognitive impairment, and vision problems. However, as in previous studies, no differences were found in the different aspects of usability based on the number of chronic diseases.

Only 38 patients out of a total of 250 who signed prior informed consent agreed to participate in this usability study. It should be noted that a mobile instant messaging app was used as the method of offering participation and there was a nonresponse rate of 80% (200/250). This means of communication, although common in current society, could have caused a lack of confidence or security in patients [40]. Among the 38 patients who agreed to participate, 6 (15.2%) decided to leave the study as a result of the stress generated by the proposed tasks or due to lack of interest. Few previous usability studies have reported the number of losses [41], perhaps due to the small number of patients involved. Wildenbos et al [22] mentioned lack of motivation as a key element to achieving acceptance of technology by older people. The benefits of using a technology should be made evident quickly and easily; otherwise, feelings of frustration and of giving up its use are likely. In a time-limited

XSL•FO
RenderX

usability evaluation, these benefits are not evident, and their nonparticipation can help to avoid feelings of uncertainty, wasting the time invested in learning a technology, or the shame of making mistakes.

Differences based on sex in the use of ICTs have been described in previous studies [37,42]. Among older people, access to technology and their degree of involvement in daily activity is greater in men than in women [37]. These differences are also identified in the different aspects of usability. In our study, the average number of errors committed per task was significantly higher in women (mean 1.4, SD 2.0) than in men (mean 0.6, SD 1.2). In addition, the proportion of women who needed personalized help to complete tasks was higher (10/15, 67%) than that in men (7/15, 47%). These differences have been largely justified by the fact that the labor participation of women has been lower, particularly in computerized jobs due to less training over the years [43].

Limitations

Although the number of participants in our study is similar to that of other studies and the findings obtained provide valuable information, a larger sample size would provide a larger data set to conduct more sophisticated and detailed statistical analyses. Moreover, given the characteristics of the research, it has not been possible to collect opinions, sensations, and emotions in relation to the technological tool that the patients experienced during task execution. For this reason, including a qualitative methodology such as focus groups [44], think-aloud tasks [45], or a user-centered cognitive walkthrough [20] could provide essential information for understanding the decision-making of patients with multimorbidity when faced with a mobile app aimed at health.

Another limitation identified is the time of tool use being limited to 48 hours. Studies such as those of Tahsin et al [46] and Baek et al [47] showed how usability results can change at different times over longer intervals of use for up to 1 year.

Conclusions

Although usability effectiveness was high, the poor efficiency and usability satisfaction results suggest that there are other factors that may interfere with these results. Sex and education level can influence the degree of completion of tasks. It has not been possible to show that multimorbidity is a key factor in the usability results of a technological tool.

Acknowledgments

This work received funding from the European Commission under the EU Horizon 2020 program (grant agreement number 875325; TeNDER project, Affective Based Integrated Care for Better Quality of Life). The paper solely reflects the views of the authors. The Commission is not responsible for the contents of this paper, or any use made thereof. Funding was also received from the semFYC private foundation to IDCG who won a 2022 doctoral thesis award, and from the Foundation for Biosanitary Research and Innovation in Primary Care (FIIBAP) for the translation and publication of this paper. The following technical partners and primary care health centers (PCHCs) and family physicians (Grupo TeNDER Atención Primaria) participated in this work: Polytechnic University of Madrid (UPM): Javier Serrano-Romero, Alberto Del Rio-Ponce, Verónica Ruiz-Bejarano; PCHC Barrio del Pilar (Madrid): Esther Labrador-Arranz, Azucena Sáez-Berlana, Elena Pérez-Gonzalez, Marta Yagüe-López; Rosa Badenes, Virtudes Enguita, Araceli Garrido-Barral; PCHC Perales del Río (Getafe): Cristina Marcos-De Frutos, Maribel León-Coronado, Victoria Ruiz-Pascual, Cesar Minué-Lorenzo, Eloína Delgado-Silván; PCHC Pedro Laín Entralgo (Alcorcón): María Esther García-Talavera, Susana Benito-Alarcón, María Luisa Muñoz-Balsa, Pablo Amich-Alemany, María Teresa Cobos-Hinojal, Lorena Córdova-Castaño, Andrea Vinssac-Rayado, Angelica Fajardo-Alcántara, Esther Alonso-Herrador, Mercedes Gea; María Mendoza-Paico, Sandra Pérez-Santos, Marta Pinedo-Hoyos, Esther Martínez-Largo, María Aránzazu Murciano-Antón, Teresa Mazarro-Enrique, Canto De Hoyos-Alonso; PCHC General Ricardos (Madrid): Esther Cabrera, Daniel Prieto-Barranco, Celeste García-Galeano, Isaías Utiel, Álvaro López, Laura Jaurena, Yolanda Rodríguez; PCHC La Rivota (Alcorcón): Inmaculada Sánchez, Esther García-Esperalta; PCHC Arroyomolinos (Arroyomolinos): Tania Beatriz Ruiz-Molina, Jesús Sánchez-Montoya, Noemí Arraña-Gallego, Susana Domínguez-Mateos; Field researchers: Rocío Acosta-Cano, Susana Marugán-Rubio, Elisa Martínez-Núñez, Laura Calderón-Jiménez.

Data Availability

The data sets generated during and/or analyzed during this study are available in the Zenodo repository [48].

Conflicts of Interest

None declared.

Multimedia Appendix 1
System Usability Scale (SUS) questionnaire.
[DOCX File , 17 KB - humanfactors_v11i1e46811_app1.docx]

Multimedia Appendix 2
Chronic diseases classified according to O'Halloran et al [33].

XSL•FO
RenderX

[DOCX File , 16 KB - humanfactors_v11i1e46811_app2.docx]

Multimedia Appendix 3
Affinity for technology interaction (ATI) scale.
[DOCX File , 16 KB - humanfactors_v11i1e46811_app3.docx]

Multimedia Appendix 4
Usability results according to the total number of patients who completed the study and total number of tasks completed with complete *P* values for all comparisons (related to Table 2).
[XLSX File (Microsoft Excel File), 12 KB - humanfactors_v11i1e46811_app4.xlsx]

Multimedia Appendix 5
Usability results by task with complete *P* values for all comparisons (related to Table 3).
[XLSX File (Microsoft Excel File), 11 KB - humanfactors_v11i1e46811_app5.xlsx]

Multimedia Appendix 6
Subgroup analysis considering patients with and without cognitive deficit.
[DOCX File , 21 KB - humanfactors_v11i1e46811_app6.docx]

References

1. Gijsen R, Hoeymans N, Schellevis FG, Ruwaard D, Satariano WA, van den Bos GA. Causes and consequences of comorbidity: a review. J Clin Epidemiol 2001 Jul;54(7):661-674. [doi: 10.1016/s0895-4356(00)00363-2] [Medline: 11438406]
2. Multimorbidity: a priority for global health research. The Academy of Medical Science. 2018 Apr. URL: https://acmedsci.ac.uk/file-download/82222577 [accessed 2022-12-03]
3. Barnett K, Mercer SW, Norbury M, Watt G, Wyke S, Guthrie B. Epidemiology of multimorbidity and implications for health care, research, and medical education: a cross-sectional study. Lancet 2012 Jul 07;380(9836):37-43 [FREE Full text] [doi: 10.1016/S0140-6736(12)60240-2] [Medline: 22579043]
4. Buetow S. Making the improbable probable: communication across models of medical practice. Health Care Anal 2014 Jun 29;22(2):160-173. [doi: 10.1007/s10728-012-0214-3] [Medline: 22743693]
5. Smith SM, Wallace E, O'Dowd T, Fortin M. Interventions for improving outcomes in patients with multimorbidity in primary care and community settings. Cochrane Database Syst Rev 2016 Mar 14;3(3):CD006560 [FREE Full text] [doi: 10.1002/14651858.CD006560.pub3] [Medline: 26976529]
6. Bodenheimer T, Lorig K, Holman H, Grumbach K. Patient self-management of chronic disease in primary care. JAMA 2002 Nov 20;288(19):2469-2475. [doi: 10.1001/jama.288.19.2469] [Medline: 12435261]
7. Leo DG, Buckley BJR, Chowdhury M, Harrison SL, Isanejad M, Lip GYH, TAILOR investigators. Interactive remote patient monitoring devices for managing chronic health conditions: systematic review and meta-analysis. J Med Internet Res 2022 Nov 03;24(11):e35508 [FREE Full text] [doi: 10.2196/35508] [Medline: 36326818]
8. Kim BY, Lee J. Smart devices for older adults managing chronic disease: a scoping review. JMIR Mhealth Uhealth 2017 May 23;5(5):e69 [FREE Full text] [doi: 10.2196/mhealth.7141] [Medline: 28536089]
9. Linn N, Goetzinger C, Regnaux J, Schmitz S, Dessenne C, Fagherazzi G, et al. Digital health interventions among people living with frailty: a scoping review. J Am Med Dir Assoc 2021 Sep;22(9):1802-1812 [FREE Full text] [doi: 10.1016/j.jamda.2021.04.012] [Medline: 34000266]
10. Jordan PW. An Introduction to Usability. London: Taylor & Francis; 1998.
11. Maramba I, Chatterjee A, Newman C. Methods of usability testing in the development of eHealth applications: a scoping review. Int J Med Inform 2019 Jun;126:95-104. [doi: 10.1016/j.ijmedinf.2019.03.018] [Medline: 31029270]
12. REPORT on the eHealth Action Plan 2012-2020 - Innovative healthcare for the 21st century. European Commission. URL: https://www.europarl.europa.eu/doceo/document/A-7-2013-0443_EN.html [accessed 2022-11-06]
13. Global Strategy on Digital Health 2020-2025. World Health Organization. URL: https://www.who.int/docs/default-source/documents/gs4dhdaa2a9f352b0445bafbc79ca799dce4d.pdf [accessed 2022-11-15]
14. Abran A, Khelifi A, Suryn W, Seffah A. Usability meanings and interpretations in ISO standards. Softw Qual J 2003;11(4):325-338. [doi: 10.1023/A:1025869312943]
15. ISO/IEC 25022:2016 Systems and software engineering. Systems and software quality requirements and evaluation (SQuaRE). Measurement of quality in use. International Organization for Standardization (ISO). 2016. URL: https://www.iso.org/standard/35746.html [accessed 2022-12-03]
16. Ergonomics of human-system interaction - Part 11: Usability: Definitions and concepts (ISO 9241-11). Tienda. 2018. URL: https://tienda.aenor.com/norma-iso-9241-11-2018-063500 [accessed 2022-10-26]

XSL•FO
RenderX

17. Saeed N, Manzoor M, Khosravi P. An exploration of usability issues in telecare monitoring systems and possible solutions: a systematic literature review. Disabil Rehabil Assist Technol 2020 Apr;15(3):271-281. [doi: 10.1080/17483107.2019.1578998] [Medline: 30794009]
18. Zapata BC, Fernández-Alemán JL, Idri A, Toval A. Empirical studies on usability of mHealth apps: a systematic literature review. J Med Syst 2015 Feb;39(2):1. [doi: 10.1007/s10916-014-0182-2] [Medline: 25600193]
19. Georgsson M, Staggers N. Quantifying usability: an evaluation of a diabetes mHealth system on effectiveness, efficiency, and satisfaction metrics with associated user characteristics. J Am Med Inform Assoc 2016 Jan;23(1):5-11 [FREE Full text] [doi: 10.1093/jamia/ocv099] [Medline: 26377990]
20. Georgsson M, Staggers N, Årsand E, Kushniruk A. Employing a user-centered cognitive walkthrough to evaluate a mHealth diabetes self-management application: A case study and beginning method validation. J Biomed Inform 2019 Mar;91:103110 [FREE Full text] [doi: 10.1016/j.jbi.2019.103110] [Medline: 30721757]
21. Ligons FM, Mello-Thoms C, Handler SM, Romagnoli KM, Hochheiser H. Assessing the impact of cognitive impairment on the usability of an electronic medication delivery unit in an assisted living population. Int J Med Inform 2014 Nov;83(11):841-848 [FREE Full text] [doi: 10.1016/j.ijmedinf.2014.07.004] [Medline: 25153770]
22. Wildenbos GA, Peute L, Jaspers M. Aging barriers influencing mobile health usability for older adults: a literature based framework (MOLD-US). Int J Med Inform 2018 Jun;114:66-75. [doi: 10.1016/j.ijmedinf.2018.03.012] [Medline: 29673606]
23. Doyle J, Murphy E, Kuiper J, Smith S, Hannigan C, Jacobs A, et al. Managing multimorbidity: identifying design requirements for a digital self-management tool to support older adults with multiple chronic conditions. 2019 Presented at: 2019 CHI Conference on Human Factors in Computing Systems; Glasgow, Scotland; May 4-9, 2019. [doi: 10.1145/3290605.3300629]
24. Chaudhry BM, Dasgupta D, Chawla N. Formative evaluation of a tablet application to support goal-oriented care in community-dwelling older adults. Proc ACM Hum Comput Interact 2022 Sep 20;6(MHCI):1-21. [doi: 10.1145/3546743]
25. Or C, Tao D. Usability study of a computer-based self-management system for older adults with chronic diseases. JMIR Res Protoc 2012 Nov 08;1(2):e13 [FREE Full text] [doi: 10.2196/resprot.2184] [Medline: 23612015]
26. Peute LW, Wildenbos G, Engelsma T, Lesselroth BJ, Lichtner V, Monkman H, et al. Overcoming challenges to inclusive user-based testing of health information technology with vulnerable older adults: recommendations from a human factors engineering expert inquiry. Yearb Med Inform 2022 Aug 02;31(1):74-81 [FREE Full text] [doi: 10.1055/s-0042-1742499] [Medline: 35654432]
27. Affective Based Integrated Care for Better quality of life (TeNDER Project). European Commission. URL: https://cordis.europa.eu/project/id/875325 [accessed 2022-11-02]
28. TeNDER. URL: https://www.tender-health.eu/project/ [accessed 2022-11-02]
29. Lozano Hernández CM, Medina-García R, de Hoyos-Alonso MC, Garrido-Barral A, Minué Lorenzo C, Sanz-Cuesta T, TeNDER Group Primary Care. Improvement in quality of life with the use of a technological system among patients with chronic disease followed up in primary care (TeNDER Project): protocol for a randomized controlled trial. JMIR Res Protoc 2023 Jul 03;12:e47331 [FREE Full text] [doi: 10.2196/47331] [Medline: 37399054]
30. Sánchez-Morillo D, Crespo M, León A, Crespo Foix LF. A novel multimodal tool for telemonitoring patients with COPD. Inform Health Soc Care 2015 Jan;40(1):1-22. [doi: 10.3109/17538157.2013.872114] [Medline: 24380372]
31. Sevilla-Gonzalez MDR, Moreno Loaeza L, Lazaro-Carrera LS, Bourguet Ramirez B, Vázquez Rodríguez A, Peralta-Pedrero ML, et al. Spanish version of the system usability scale for the assessment of electronic tools: development and validation. JMIR Hum Factors 2020 Dec 16;7(4):e21161 [FREE Full text] [doi: 10.2196/21161] [Medline: 33325828]
32. Bangor A, Kortum PT, Miller JT. An empirical evaluation of the System Usability Scale. Int J Hum Comput Interact 2008 Jul 30;24(6):574-594. [doi: 10.1080/10447310802205776]
33. O'Halloran J, Miller G, Britt H. Defining chronic conditions for primary care with ICPC-2. Fam Pract 2004 Aug;21(4):381-386. [doi: 10.1093/fampra/cmh407] [Medline: 15249526]
34. Franke T, Attig C, Wessel D. A personal resource for technology interaction: development and validation of the Affinity for Technology Interaction (ATI) scale. Int J Hum Comput Interact 2018 Mar 30;35(6):456-467. [doi: 10.1080/10447318.2018.1456150]
35. Harris PA, Taylor R, Minor BL, Elliott V, Fernandez M, O'Neal L, REDCap Consortium. The REDCap consortium: building an international community of software platform partners. J Biomed Inform 2019 Jul;95:103208 [FREE Full text] [doi: 10.1016/j.jbi.2019.103208] [Medline: 31078660]
36. Harris PA, Taylor R, Thielke R, Payne J, Gonzalez N, Conde JG. Research electronic data capture (REDCap)--a metadata-driven methodology and workflow process for providing translational research informatics support. J Biomed Inform 2009 Apr;42(2):377-381 [FREE Full text] [doi: 10.1016/j.jbi.2008.08.010] [Medline: 18929686]
37. Kim J, Lee HY, Christensen MC, Merighi JR. Technology access and use, and their associations with social engagement among older adults: do women and men differ? J Gerontol B Psychol Sci Soc Sci 2017 Sep 01;72(5):836-845. [doi: 10.1093/geronb/gbw123] [Medline: 28073816]
38. Smith A. Older adults and technology use. Pew Research Center. 2014 Apr. URL: http://www.pewinternet.org/2014/04/03/older-adults-and-technology-use/ [accessed 2022-11-02]

XSL•FO
RenderX

39. Wildenbos GA, Jaspers MWM, Schijven MP, Dusseljee-Peute LW. Mobile health for older adult patients: using an aging barriers framework to classify usability problems. Int J Med Inform 2019 Apr;124:68-77 [FREE Full text] [doi: 10.1016/j.ijmedinf.2019.01.006] [Medline: 30784429]
40. Kiat BW, Chen W. Mobile instant messaging for the elderly. Proced Comput Sci 2015;67:28-37. [doi: 10.1016/j.procs.2015.09.246]
41. Czaja SJ, Sharit J, Lee CC, Nair SN, Hernández MA, Arana N, et al. Factors influencing use of an e-health website in a community sample of older adults. J Am Med Inform Assoc 2013 Mar 01;20(2):277-284 [FREE Full text] [doi: 10.1136/amiajnl-2012-000876] [Medline: 22802269]
42. Milagros S, Lidia A, Cecilia C. Mujeres y digitalización. De las brechas a los algoritmos. Instituto de la Mujer y para la Igualdad de Oportunidades. Ministerio de Igualdad. 2020. URL: https://www.inmujeres.gob.es/diseno/novedades/M_MUJERES_Y_DIGITALIZACION_DE_LAS_BRECHAS_A_LOS_ALGORITMOS_04.pdf [accessed 2022-11-02]
43. Hargittai E, Dobransky K. Old dogs, new clicks: digital inequality in skills and uses among older adults. Can J Commun 2017 May 23;42(2):195-212. [doi: 10.22230/cjc.2017v42n2a3176]
44. Brown W, Yen P, Rojas M, Schnall R. Assessment of the Health IT Usability Evaluation Model (Health-ITUEM) for evaluating mobile health (mHealth) technology. J Biomed Inform 2013 Dec;46(6):1080-1087 [FREE Full text] [doi: 10.1016/j.jbi.2013.08.001] [Medline: 23973872]
45. Bolle S, Romijn G, Smets EMA, Loos EF, Kunneman M, van Weert JCM. Older cancer patients' user experiences with web-based health information tools: a think-aloud study. J Med Internet Res 2016 Jul 25;18(7):e208 [FREE Full text] [doi: 10.2196/jmir.5618] [Medline: 27457709]
46. Tahsin F, Tracy S, Chau E, Harvey S, Loganathan M, McKinstry B, et al. Exploring the relationship between the usability of a goal-oriented mobile health application and non-usage attrition in patients with multimorbidity: a blended data analysis approach. Digit Health 2021 Oct 05;7:20552076211045579 [FREE Full text] [doi: 10.1177/20552076211045579] [Medline: 34868614]
47. Baek JY, Na SH, Lee H, Jung H, Lee E, Jo M, et al. Implementation of an integrated home internet of things system for vulnerable older adults using a frailty-centered approach. Sci Rep 2022 Feb 04;12(1):1922. [doi: 10.1038/s41598-022-05963-9] [Medline: 35121795]
48. Medina-García R, Lopez-Rodríguez JA. A technological tool aimed at self-care in patients with multimorbidity: cross-sectional usability study (complete study data). Zenodo. URL: https://zenodo.org/records/10848551 [accessed 2024-03-21]

Abbreviations

ATI: affinity for technology interaction
DCL: data collection logbook
ICT: Information and communication technology
ISO: International Organization for Standardization
MMSE: Mini-Mental State Evaluation
REDCap: Research Electronic Data Capture
SUS: System Usability Scale
TeNDER: Affective Based Integrated Care for Better Quality of Life

Edited by Y Senathirajah; submitted 26.02.23; peer-reviewed by N Mungoli, B Chaudhry; comments to author 22.08.23; revised version received 05.09.23; accepted 28.02.24; published 05.04.24.

Please cite as:
Medina-García R, López-Rodríguez JA, Lozano-Hernández CM, Ruiz Bejerano V, Criscio P, Del Cura-González I, TeNDER Atención Primaria Group
A Technological Tool Aimed at Self-Care in Patients With Multimorbidity: Cross-Sectional Usability Study
JMIR Hum Factors 2024;11:e46811
URL: https://humanfactors.jmir.org/2024/1/e46811
doi:10.2196/46811
PMID:38578675

©Rodrigo Medina-García, Juan A López-Rodríguez, Cristina María Lozano-Hernández, Verónica Ruiz Bejerano, Paride Criscio, Isabel Del Cura-González, TeNDER Atención Primaria Group. Originally published in JMIR Human Factors (https://humanfactors.jmir.org), 05.04.2024. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic

XSL•FO
RenderX

information, a link to the original publication on https://humanfactors.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Multimodal In-Vehicle Hypoglycemia Warning for Drivers With Type 1 Diabetes: Design and Evaluation in Simulated and Real-World Driving

Caterina Bérubé[1], PhD; Martin Maritsch[1], PhD; Vera Franziska Lehmann[1,2], Dr med, PhD; Mathias Kraus[3], Prof Dr; Stefan Feuerriegel[1,4], Prof Dr; Thomas Züger[1,2,5], Dr med, PhD; Felix Wortmann[1,6], Prof Dr; Christoph Stettler[1,2], Prof Dr; Elgar Fleisch[1,6], Prof Dr; A Baki Kocaballi[7], PhD; Tobias Kowatsch[1,6,8], Prof Dr

[1]Centre for Digital Health Interventions, Department of Management, Technology and Economics, ETH Zurich, Zurich, Switzerland
[2]Department of Diabetes, Endocrinology, Nutritional Medicine and Metabolism, Bern University Hospital, University of Bern, Bern, Switzerland
[3]School of Business, Economics and Society, Friedrich-Alexander-Universit¨at Erlangen-Nürnberg, Nürnberg, Germany
[4]LMU Munich School of Management, LMU Munich, Munich, Germany
[5]Department of Endocrinology and Metabolic Diseases, Kantonsspital Olten, Olten, Switzerland
[6]Institute of Technology Management, University of St.Gallen, St Gallen, Switzerland
[7]School of Computer Science, University of Technology Sydney, Sydney, Australia
[8]Institute for Implementation Science in Health Care, University of Zurich, Zurich, Switzerland

Corresponding Author:
Tobias Kowatsch, Prof Dr
Institute of Technology Management
University of St.Gallen
Office 62-314
St.Jakob-Strasse 21
St Gallen, 9000
Switzerland
Phone: 41 71 224 7244
Email: tobias.kowatsch@unisg.ch

Abstract

Background: Hypoglycemia threatens cognitive function and driving safety. Previous research investigated in-vehicle voice assistants as hypoglycemia warnings. However, they could startle drivers. To address this, we combine voice warnings with ambient LEDs.

Objective: The study assesses the effect of in-vehicle multimodal warning on emotional reaction and technology acceptance among drivers with type 1 diabetes.

Methods: Two studies were conducted, one in simulated driving and the other in real-world driving. A quasi-experimental design included 2 independent variables (blood glucose phase and warning modality) and 1 main dependent variable (emotional reaction). Blood glucose was manipulated via intravenous catheters, and warning modality was manipulated by combining a tablet voice warning app and LEDs. Emotional reaction was measured physiologically via skin conductance response and subjectively with the Affective Slider and tested with a mixed-effect linear model. Secondary outcomes included self-reported technology acceptance. Participants were recruited from Bern University Hospital, Switzerland.

Results: The simulated and real-world driving studies involved 9 and 10 participants with type 1 diabetes, respectively. Both studies showed significant results in self-reported emotional reactions ($P<.001$). In simulated driving, neither warning modality nor blood glucose phase significantly affected self-reported arousal, but in real-world driving, both did ($F_{2,68}=4.3$; $P<.05$ and $F_{2,76}=4.1$; $P=.03$). Warning modality affected self-reported valence in simulated driving ($F_{2,68}=3.9$; $P<.05$), while blood glucose phase affected it in real-world driving ($F_{2,76}=9.3$; $P<.001$). Skin conductance response did not yield significant results neither in the simulated driving study (modality: $F_{2,68}=2.46$; $P=.09$, blood glucose phase: $F_{2,68}=0.3$; $P=.74$), nor in the real-world driving study (modality: $F_{2,76}=0.8$; $P=.47$, blood glucose phase: $F_{2,76}=0.7$; $P=.5$). In both simulated and real-world driving studies, the voice+LED warning modality was the most effective (simulated: mean 3.38, SD 1.06 and real-world: mean 3.5, SD 0.71) and urgent (simulated: mean 3.12, SD 0.64 and real-world: mean 3.6, SD 0.52). Annoyance varied across settings. The standard

XSL•FO
RenderX

warning modality was the least effective (simulated: mean 2.25, SD 1.16 and real-world: mean 3.3, SD 1.06) and urgent (simulated: mean 1.88, SD 1.55 and real-world: mean 2.6, SD 1.26) and the most annoying (simulated: mean 2.25, SD 1.16 and real-world: mean 1.7, SD 0.95). In terms of preference, the voice warning modality outperformed the standard warning modality. In simulated driving, the voice+LED warning modality (mean rank 1.5, SD rank 0.82) was preferred over the voice (mean rank 2.2, SD rank 0.6) and standard (mean rank 2.4, SD rank 0.81) warning modalities, while in real-world driving, the voice+LED and voice warning modalities were equally preferred (mean rank 1.8, SD rank 0.79) to the standard warning modality (mean rank 2.4, SD rank 0.84).

Conclusions: Despite the mixed results, this paper highlights the potential of implementing voice assistant–based health warnings in cars and advocates for multimodal alerts to enhance hypoglycemia management while driving.

Trial Registration: ClinicalTrials.gov NCT05183191; https://classic.clinicaltrials.gov/ct2/show/NCT05183191, ClinicalTrials.gov NCT05308095; https://classic.clinicaltrials.gov/ct2/show/NCT05308095

(JMIR Hum Factors 2024;11:e46967) doi:10.2196/46967

KEYWORDS

digital health; voice assistant; ambient lighting; in-vehicle technology; health state; diabetes; hypoglycemia; warning; emotional reaction; technology acceptance; mobile phone; diabetes; implementation

Introduction

Overview

Around 9 million adults worldwide experience type 1 diabetes mellitus (T1DM) [1]. One of the most relevant acute complications associated with T1DM is hypoglycemia (ie, low blood glucose). This condition is associated with impaired cognitive, executive, and psychomotor function [2-4] and is linked to driving mishaps [5-7].

Previous work introduced the development of a voice warning for hypoglycemia while behind the wheel, whereas the voice assistant (VA) would work as a warning interface [8]. The hypoglycemia warning was intended as an app compatible with the VA that is already available in the car and that would allow delivering an alert in a hands-free manner. The study reported on the iterative development and evaluation of an in-vehicle hypoglycemia voice warning. It demonstrated that it is deemed useful and effective by drivers with T1DM, especially if the warning is kept simple and direct (ie, avoiding initiating a conversation with the driver). However, the paper did not investigate the effect of proactive behavior in the VA in such a context. Proactivity in VAs can cause a startling reaction, which is prone to annoyance [9] and driving impairments [10-12].

Ambient lighting can communicate with drivers without distracting them from their main task [13] or eliciting a strong emotional response [14]. This interface has been investigated as an indicator of several driving-relevant events, such as obstacle warnings or vehicle-state communication [15]. However, to the best of our knowledge, in-vehicle ambient lighting through LEDs has never been investigated as an indicator of a critical health state.

Our previous work [8] tested the concept of an in-vehicle voice warning delivered by the VA with healthy participants and then with individuals with T1DM, both in a driving simulator and a real car. The concept was developed following an iterative approach, and study participants provided feedback that we used to enhance the voice warning and test it on new participants. Thus, our previous work focused on technology acceptance and on improving it through user feedback. However, the voice warning was not evaluated against a standard warning (ie, beep with text), which could be used as a benchmark. Moreover, our previous work did not focus on the emotional reaction generated by getting a warning while driving. Therefore, this work investigates the effect of LEDs (ie, a possible solution to alleviate emotional reaction) and is to be understood as a continuation of our previous work. To foster experimental control and external validity, the same procedure is replicated in a simulated driving setting (ie, a computer simulator) and a real-world driving setting (ie, a car in a closed circuit).

Background

Hypoglycemia Warnings

Hypoglycemia is a common complication of diabetes. The monitoring of blood glucose is essential to prevent hypoglycemia. Intermittent self-monitoring of blood glucose, flash glucose monitoring, and continuous glucose monitoring are commonly used methods. However, these methods are not adapted to the in-vehicle context, as they require the driver to visually attend to a handheld mobile device displaying the current blood glucose value. This behavior is known to impair driving performance [16], thus leading to dangerous situations while driving.

Tentative hands-free solutions have been proposed to address this issue in academia [17] and in the community of individuals with T1DM [18]. Specifically, prior research [17] suggested using vehicles as a platform to display blood glucose data on infotainment screens. Moreover, a digital community [18] created an open-source program to show their continuous glucose monitoring data on infotainment screens. However, these solutions are limited to visual information display, thus failing to be ergonomically suitable for the in-vehicle context while driving. Therefore, solutions must be developed, which can provide hypoglycemia warnings while driving. One approach is to use voice-first warnings (ie, delivered by the built-in in-vehicle VA), where the driver can be informed of the issue without having to attend to a display.

XSL•FO
RenderX

In-Vehicle Warnings

In-vehicle health-state warning systems are a part of advanced driver–assistance systems [19]. From a human-computer interaction perspective, in-vehicle warnings should be effective and communicate urgency without being annoying [20]. Currently, in-vehicle warnings vary from classic car warnings, visually presented on the dashboard with traffic-light colors and unspecific tones, to advanced driver–assistance system warnings that use visual, auditory, and haptic modalities [21]. Even though the visual signal should be redundant to the auditory and haptic signals, some driver-state warnings, such as the driver attention alert, are predominantly visual (eg, mug symbol with an indicative text such as "Time for a break").

To decrease the demand for drivers' visual attention, it is necessary to develop attention-attractive warnings without relying on visual displays as the main source of information. One approach would be to use the in-vehicle VA already built into the car to warn or alert drivers for critical situations, such as hypoglycemia (or drowsiness). This approach could ensure warnings' effectiveness while reducing drivers' visual distraction.

In-Vehicle VAs

VAs are increasingly being integrated into cars [22-24], allowing digital health interventions to be delivered via such an interface in a scalable way. In the in-vehicle context, VAs have ergonomic and experiential advantages: they reduce visual distraction compared to other infotainment technologies [25], foster a natural interaction [26], create a sense of social presence [27], and increase engagement [28]. Ultimately, VAs can create a sense of being in the presence of a copilot [29]. Therefore, in-vehicle VAs have great potential for delivering real-time and effective hypoglycemia warnings to drivers while driving.

Proactive VAs and the Risk of Startle

Proactivity is not part of the current common mental model of a VA. However, it does not necessarily affect driving performance [30] and is well-accepted by drivers [31,32]. Nevertheless, a sudden auditory stimulus can create a startling reaction, which could interfere with driving performance [10,33]. Hence, when it comes to critical situations such as hypoglycemia [34], it is important to develop warnings that gradually prepare the driver to be receptive to them. Ambient lighting can be used to gradually prepare the driver and add information without consequentially distracting the driver [13]. This technology has been previously investigated for in-vehicle driving behavior support such as collision and blind warnings, lane change decision support, and speed and attention direction recommendations [15]. In addition, it has been investigated to inform the driver about the vehicle's decision-making in autonomous cars [35].

Objectives

To this end, a hypoglycemia warning delivered by a mock-up of a built-in in-vehicle VA is designed and tested with individuals with T1DM and compared to a standard format of in-car warning (ie, unspecific alert tone with visual information) regarding driver experience.

Specifically, this work aims (1) to design a hands-free multimodal health intervention for hypoglycemia (ie, warning) compatible with the in-vehicle context and (2) to investigate the effect of warning modality (visual, vocal, and vocal with ambient lighting) on the emotional reaction and the acceptance of such technology.

Methods

Overview

Two studies were carried out, a simulated driving study and a real-world driving study. Across these studies, participant recruitment, study design, material and apparatus, procedure, and data analysis were the same. The difference lied in the setting. For this reason, all the following subsections, except for *Setting*, are described only once.

Participants

Sampling, Inclusion Criteria, and Compensation

Patients diagnosed with T1DM attending the diabetes outpatient clinic of the Bern University Hospital were recruited. A physician (VFL) of the study team performed recruitment during regular outpatient visits with a face-to-face assessment. For the simulated driving study, participants were recruited between November 2021 and March 2022. For the real-world driving study, participants were recruited between April and June 2022. Inclusion criteria were age between 21 and 60 years, hemoglobin A_{1c}≤9% (ie, a blood test indicating how well the patient's diabetes is being controlled), functional insulin treatment (with insulin pump therapy or multiple daily injections) for at least 3 months with good knowledge of insulin self-management, possession of a Swiss driver's license at least 3 years before study inclusion, and have driven at least once in the last 6 months. Each participant received an expense allowance of US $209.62 to cover general expenses caused by study participation (eg, transport).

Experience and Beliefs Questionnaires

Upon inclusion in the study, participants were asked to report the frequency of driving per week, their previous use of in-vehicle VAs, and their technology affinity. Technology affinity was assessed with the 16-item Technology Readiness Index [36]. This scale measures constructs susceptible to influencing the adoption of cutting-edge technology, such as optimism, innovativeness, discomfort, and insecurity.

Study Design

The study was designed as quasi-experimental with 2 independent variables, that is, blood glucose phase and warning modality, and 1 main dependent variable, that is, emotional reaction. The blood glucose variable had 3 levels, that is, euglycemia, decreasing, and hypoglycemia (see the *Procedure* section). The warning modality variable had 3 levels as well, that is, standard, voice, and voice+LED (see the *Warning* section). The blood glucose phase was varied in a nonrandomized fashion (see the *Procedure* section), while the warning modality was pseudorandomized, and each modality was crossed with each blood glucose phase. Secondary outcomes included self-reported user experience measures, such as

XSL•FO
RenderX

warning acceptance, perceived urgency, alerting effectiveness, annoyance, and preference.

Material and Apparatus

Overview

In this section, the operationalization of the design variables is described. An overview is listed in Table 1.

Table 1. Study design variables.

Variable	Tool	Levels or values
Blood glucose[a]	Controlled hypoglycemia protocol [37]	Normal, decreasing, and hypoglycemia
Warning modality[a]	Hypoglycemia warning app [38]	Standard, voice, and voice+LED
Emotional reaction (objective)[b]	Empatica E4	Skin conductance response
Emotional reaction (subjective)[b]	Affective Slider [39]	Score (0-100)
Warning perceived urgency[b]	Baldwin and Moore scale [20]	Score (1=very urgent and 5=very insignificant)
Warning alerting effectiveness[b]	Baldwin and Moore scale [20]	Score (1=effective and 5=ineffective)
Warning annoyance[b]	Baldwin and Moore scale [20]	Score (1=I dislike it very much and 5=I like it very much)
Warning acceptance[b]	Van Der Laan Acceptance Scale [40]	Score (–2=negative extreme and +2=positive extreme)
Warning preference[b]	Arbitrary 3-point scale	Rank (1=best and 3=worst)

[a]Manipulation.
[b]Measure.

Blood Glucose

Blood glucose was manipulated by inserting 2 intravenous catheters: one for blood glucose measurement with an interval of 5-10 minutes and the other for the infusion of a combination of insulin and glucose, according to the patient's current blood glucose and the experimental target blood glucose range. Euglycemia (ie, normal blood glucose) was defined as a concentration of 5-8 mmol/L; decreasing blood glucose was identified when blood glucose was below the euglycemia range (5-8 mmol/L) and progressing toward a target hypoglycemic range (3-3.5 mmol/L); and hypoglycemia was defined as a concentration of 3-3.5 mmol/L. For more technical details, refer to related research [37].

Warning

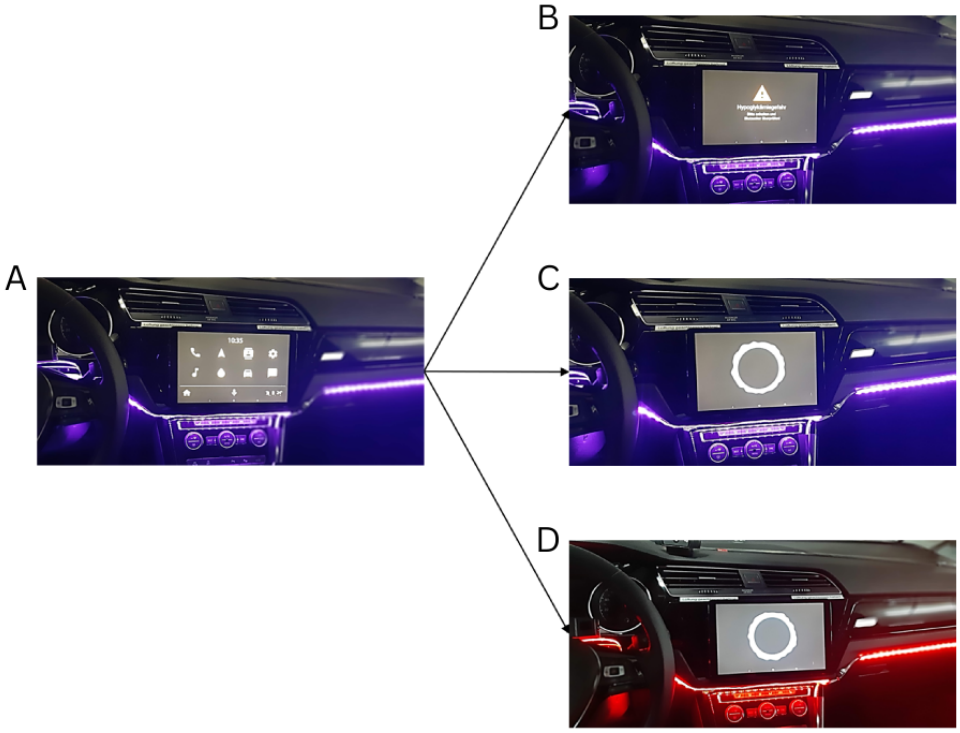
The Hypoglycemia warning app operationalized into a tablet (model SM-T590, Samsung) simulated the infotainment system's touchscreen, and LED strips (RGB Light Strip Pro, Cololight) simulated the interior ambient lighting. The warning system was controlled through the Wizard of Oz method [41], that is, the system was controlled by the experimenter behind the scene and acting as if it was fully automated.

The LED strips had 60 LEDs per meter and could be remotely controlled with the Cololight app (Klaus Stephan GmbH). The tablet was used to run the Hypoglycemia warning app (publicly available on GitHub [38]), remotely controlled via a remote desktop app (AnyDesk, AnyDesk Software GmbH).

The warning system had 4 possible states: a default state and 3 intervention modalities (standard, voice, and voice+LED). The default state involved the LED strip being turned on in blue and the tablet showing a fake navigation menu (Figure 1A). The standard modality displayed a yellow warning sign with an informative text and was accompanied by an earcon. The text said "Risk of hypoglycemia. Please pull over and verify blood sugar." The LED strips remained blue (Figure 1B). The voice modality displayed a VA animation accompanied by a prerecorded synthesized female voice (de-DE-Wavenet-C with speed=0.85 and pitch=–3.20, Google Inc). The voice said "I have detected a risk of hypoglycemia. Please pull over safely and verify your blood sugar" (translation from the German formulation "*Ich habe eine Hypogefahr erkannt. Bitte sicher anhalten und deinen Blutzucker überprüfen*"). The voice warning was designed based on the results reported in our previous work [8]. Once again, the LED strips remained blue (Figure 1C). The voice+LED modality displayed the same VA animation and prerecorded synthesized female voice but, before the onset of the voice warning, the LED strips turned red (Figure 1D).

XSL•FO
RenderX

Figure 1. Illustration of the warning app and modalities. (A) The default state was shown during the drive and simulated the infotainment menu of the car. The (B) standard, (C) voice, and (D) voice+LED modalities were activated when a warning was delivered.



Participants knew that they would receive a warning during each drive. Still, the warning presentation was pseudorandomized, where they would receive a complete permutation of the 3 warning modalities within a blood glucose phase. Participants were not explicitly informed about which warning was "the intervention of interest" and which one was the "comparator."

Objective Emotional Reaction

Emotional reaction was measured physiologically through skin conductance response (SCR). SCR is the result of the sympathetic nervous system promptly regulating the activity of the sweat glands in response to a stimulus. This measure is associated with emotional arousal [42] and can be used to measure event-related emotional reactions objectively [43]. In this study, the Empatica E4 (Empatica Inc), a Conformité Européenne–certified wristband collecting physiological data in real time, was used. Participants wore the E4 during the main visit (see the *Procedure* section). Note that this measure provided solely the arousal dimension of emotion.

Subjective Emotional Reaction

Emotional reaction was also measured subjectively through the Affective Slider [39]. This digital scale is a self-reporting tool measuring valence and arousal on 2 separate sliders. Participants did not see any numerical anchor, but the score ranged from 0 to 100. Thus, valence is rated between a frowning and a smiling face (0=frowning and 100=smiling), and arousal between a sleepy and a widely awake face (0=sleepy and 100=widely awake).

Warning Perceived Urgency, Alerting Effectiveness, and Annoyance

To measure the perceived urgency mapping of the 3 modalities, for each modality, participants rated the perceived urgency, alerting effectiveness, and annoyance according to a scale from prior work [20]. The 3 dimensions were rated using a 5-point Likert scale (1=very urgent or effective or I like it very much and 5=very insignificant or ineffective or I dislike it very much). This questionnaire was filled out during the posttest visit (see the *Procedure* section).

Warning Acceptance

To compare the acceptance of the 3 modalities, participants filled out the Van Der Laan Acceptance Scale [40], once per modality. This scale consists of the 2 constructs, usefulness and satisfaction, with items answered on a 5-point semantic differential from –2 to +2, which means participants had to select a point between 2 opposite adjectives (eg, unpleasant or pleasant). This questionnaire was filled out during the posttest visit (see the *Procedure* section).

Warning Preference

To formalize their preference, patients were asked to rank the 3 modalities from best to worst. The scale was implemented as a radio button questionnaire with 1 item per modality (ie, beep with a warning sign and text, voice, and LED with voice) and a 3-point scale (ie, 1=best and 3=worst). Participants were also encouraged to provide comments to their answers, which were topically (ie, without verbatim transcription) recorded in written form. This questionnaire was filled out during the posttest visit (see the *Procedure* section).

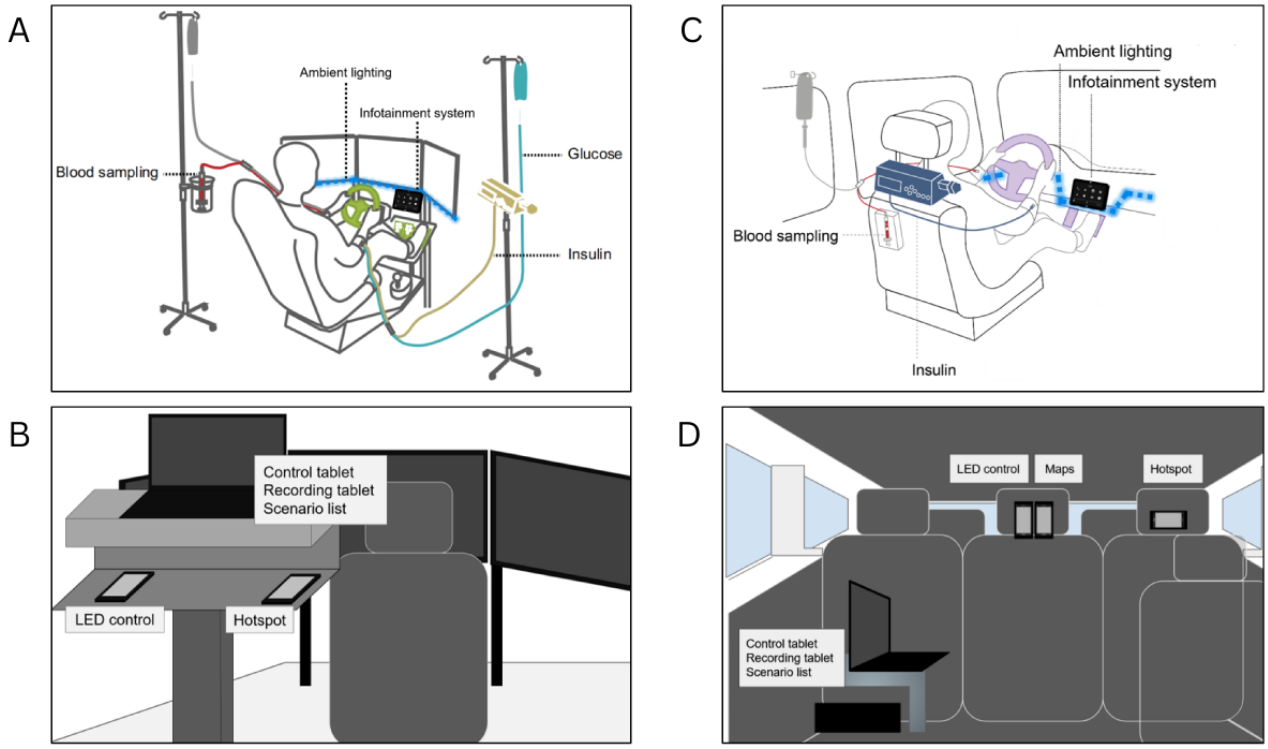
XSL•FO
RenderX

Setting

Simulated Driving

Patients used a driving simulator (Carnetsoft Inc), featuring 3 monitors displaying the front, left, and right views of the driving cabin. The central monitor also displayed the cockpit and navigation arrows, which directed the patient through the environment. To control the simulator, participants used a steering wheel and pedals (Logitech Driving Force G29, Logitech), set to the automatic transmission. The simulator was connected to a stereo speaker and maintained at a constant volume (Figure 2A and C). The infotainment system simulator tablet (see the *Warning* section) was placed under the right side of the central monitor and connected to the simulator's sound system. The LED lights were attached at the bottom of the 3 monitors. Figure 2A and 2C illustrates the patient's setup.

Figure 2. Comparison of patient's and experimenter's setup between simulated and real-world driving. The left figures represent the simulated driving setting, the right figures represent the real-world driving setting, the top figures (A and C) show the patient's setup, and the bottom figures (B and D) show the experimenter's setup.



The experimenter was standing behind the patient and controlled the LED stripes with a smartphone (Redmi Note, Xiaomi Inc) and the tablet via a laptop computer (ThinkPad X1 Carbon, Lenovo PC HK Ltd). A stopwatch app was used to manually onset the warning. Figure 2B and 2D illustrates the experimenter's setup.

Three environments were used, namely, highway, countryside, and town, with the highway being the easiest to navigate due to variable traffic but no turns, the countryside having a moderated amount of traffic and turns, and the town being the most difficult with the most turns and traffic. Participants drove in the environments for about 5 minutes before receiving a hypoglycemia warning (run-in phase).

Real-World Driving

Patients drove in a minivan (Touran, Volkswagen) with an automatic transmission. The car was equipped with dual pedals to allow for intervention from a trained driving instructor, in case of emergency. In case the instructor needed to intervene, the event was recorded.

The infotainment system simulator tablet (see the *Warning* section) was placed on top of the infotainment screen and connected to the car's sound system, maintained at a constant volume. The LED lights were attached along the cockpit from the left to the right extremities, passing by under the steering wheel, the infotainment system, and above the aperture of the glove compartment. The experimenter was sitting in the third row of the car and controlled the LED stripes with a smartphone and the tablet via a laptop computer (6th Gen ThinkPad X1 Carbon, Lenovo PC HK Ltd). A Google Map was used to manually onset the warning.

Patients were exposed to real-world driving on a test track provided by the Swiss Federal Department of Defense, Civil Protection and Sports. The driving scenarios on the track were designed to correspond with simulated environments used in the simulator setting (ie, highway, countryside, and town), featuring various driving elements such as turns, crossroads, stop signs, and a pedestrian crossing equipped with a dummy. As traffic simulation was not feasible, artificial obstacles, including boxes and traffic pylons, were used. Participants drove in the environments for 5-7 minutes before receiving a hypoglycemia warning (run-in phase).

Ethical Considerations

The experiments were approved within the context of this project by the cantonal ethics commission of Bern, Switzerland

XSL•FO
RenderX

(BASEC2020-00685 and BASEC2021-02381). Before any study-related procedure, informed consent specifying the analysis and the study protocol presented in this paper was obtained in written form from all participants. All collected data were deidentified by associating individual data to a numerical identification number. The data reported in this paper are part of the HEADWIND Study, a clinical trial registered under ClinicalTrials.gov (Part 3: NCT05183191 and Part 4: NCT05308095).

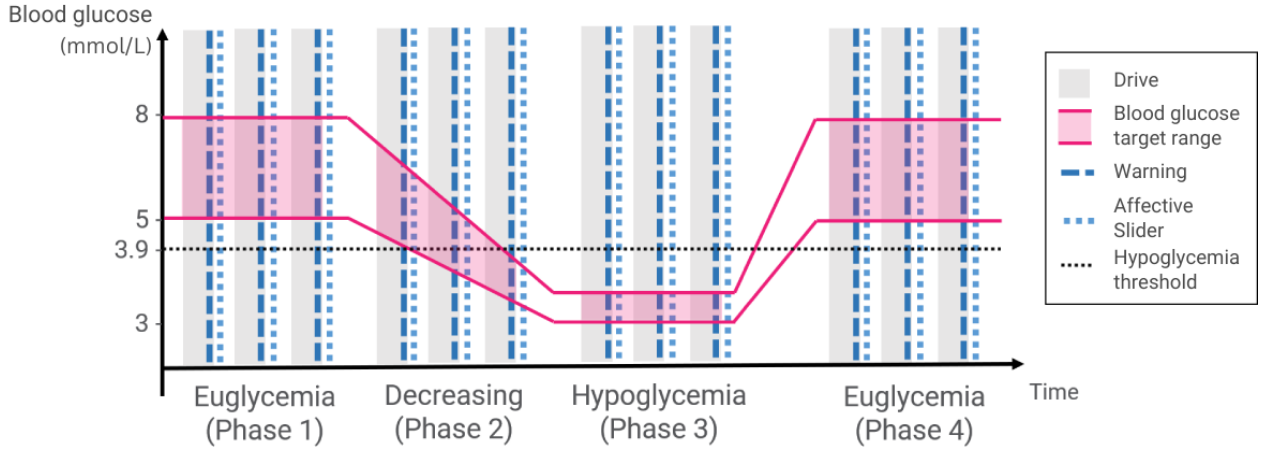
Procedure

The procedure was divided into 3 visits. In a pretest visit, patients were welcomed to the Bern University Hospital, informed about the procedure and the warnings, and asked to fill out demographic and experience and beliefs questionnaires.

In the main visit, participants were welcomed to the relevant setting for the blood glucose manipulation and the objective emotional reaction measurements. Participants were aware that their blood sugar would be manipulated to reach specific goal ranges, but they were not informed of their blood glucose during the experiment. In the real-world driving setting, the driving instructor was also aware of the blood glucose manipulation and was blinded to the current blood glucose. The experimental team was not blinded to the blood glucose level.

Each participant went through a fixed sequence of blood glucose phases: (1) a first drive with normal blood glucose (phase 1: euglycemia), where participants were first experiencing all types of environment and warning (and was thus considered as training); (2) a phase where blood glucose was progressively decreasing toward the moderate hypoglycemia threshold (ie, <3.0 mmol/L) with a target range of 3-3.5 mmol/L (phase 2: decreasing); (3) a phase with stable moderate hypoglycemia (phase 3: hypoglycemia); and (4) a final phase with normal blood glucose (phase 4: euglycemia). A warning was delivered at the end of each drive to explore the effect of the blood glucose phase. Participants drove in the 3 types of environments in each phase. The sequence of environment type was pseudorandomized [8], that is, participants were exposed to all 3 warning modalities within each phase, but the sequence of modalities within 1 phase was random. Similarly, the warning modality was pseudorandomized to balance modality with environments. Once participants received a warning, they were expected to stop, and the drive came to an end. At the end of each drive, participants filled out the Affective Slider referring to the warning they just received. Figure 3 shows an overview of the procedure.

Figure 3. Overview of the procedure. The vertical gray bars represent the drives, the ribbon delimited by solid lines represents the blood glucose manipulation, the vertical dashed lines represent the warning deliveries, the vertical dotted lines represent the Affective Slider submission, and the horizontal dotted line represents the hypoglycemia threshold.



In a posttest visit, participants were once more exposed to the 3 warning modalities and were required, after each exposure, to fill the Baldwin and Moore scale [20] and the Van Der Laan Acceptance Scale. Finally, they ranked the 3 warning modalities.

Data Analysis

The continuous variables of sample characteristics (ie, demographics and previous experience) are presented with mean and SD. Frequency variables of sample characteristics are presented in count numbers (ie, n) and percentages of the total experiment sample.

Emotional reaction (objective and subjective) measures were analyzed as a function of blood glucose (excluding phase 1) and warning modality and verified with a mixed-effects linear model, ANOVA test, and a significance threshold of P=.05. Effect size was calculated with partial η^2 (0.01 indicates a small effect, 0.06 indicates a medium effect, and 0.14 indicates a large effect). Moreover, the objective emotional reaction was analyzed following established guidelines [43]: SCR (ie, rapid phasic component) was standardized for individual differences by dividing the SCR signal by the individual maximum SCR and by reducing the noise. In addition, SCR was calculated by considering the change in skin conductance between the average skin conductance in the 5-second window before the warning onset and the average skin conductance in the 5-second window after the warning onset itself (including latency of 1 second). For each measure of emotional reaction, a mixed-effects linear model was estimated with warning modality (3 levels: standard, voice, and voice+LED) and blood glucose (3 levels: normal, decreasing, and hypoglycemia) as independent variables and

XSL•FO
RenderX

with emotional reaction measure (ie, either self-reported arousal, self-reported valence, or SCR) as the dependent variable.

Warning evaluations (ie, perceived urgency, alerting effectiveness, annoyance, and acceptance) were aggregated with means and SDs, presented in a table. Moreover, perceived urgency, alerting effectiveness, and annoyance were centered on 2 to match the acceptance scores, for the sake of comparison.

Preference ranking was aggregated across modalities in the frequency of rank as best, middle, or worst (ie, how many times 1 modality was ranked as the best, middle, or worst in comparison to the other 2). If participants were commenting on their choices, highlight note-taking was performed.

Data analysis and graphical representations were performed using RStudio (Posit Software) packages such as *lmerTest* or mixed-effects linear modeling and *ggplot2* and patchwork for data visualization. All results are separated by experiment (ie, simulated vs real-world driving) and juxtaposed to allow direct comparison.

Results

Overview

The data of 2 participants of the simulated driving study were excluded due to partial data loss. In the real-world driving setting, the driving instruction had to intervene in 1 instance, as the participants did not follow the driving path (ie, did not turn left) during phase 4 (ie, while in euglycemia).

Sample Characteristics

Overall, the majority were male participants, who drove multiple times per week and did not have previous experience with in-vehicle VAs. Participants were approximately 40 years of age and had a Technology Readiness Index between 3.5 and 4 (over a maximum of 5). Details are shown in Table 2.

Table 2. Sample characteristics across studies.

Characteristics	Simulated (n=9)	Real-world (n=10)
Sex, n (%)		
Male	8 (89)	7 (67)
Female	1 (11)	3 (33)
Age (years), mean (SD)	45.7 (11.79)	37.3 (11.1)
Frequency of driving, n (%)		
1 time per month	1 (11)	1 (10)
2-5 times per month	__[a]	2 (20)
2-5 times per week	5 (56)	4 (40)
Every day	3 (33)	3 (30)
Previous use of in-vehicle voice assistants, n (%)		
Never	4 (44)	6 (60)
Rarely	1 (11)	2 (20)
Sometimes	2 (22)	1 (10)
Often	2 (22)	1 (10)
TRI[b] (over a maximum of 5), mean (SD)	3.6 (0.62)	3.9 (0.5)

[a]Not available.

[b]TRI: Technology Readiness Index.

Emotional Reaction

Overview

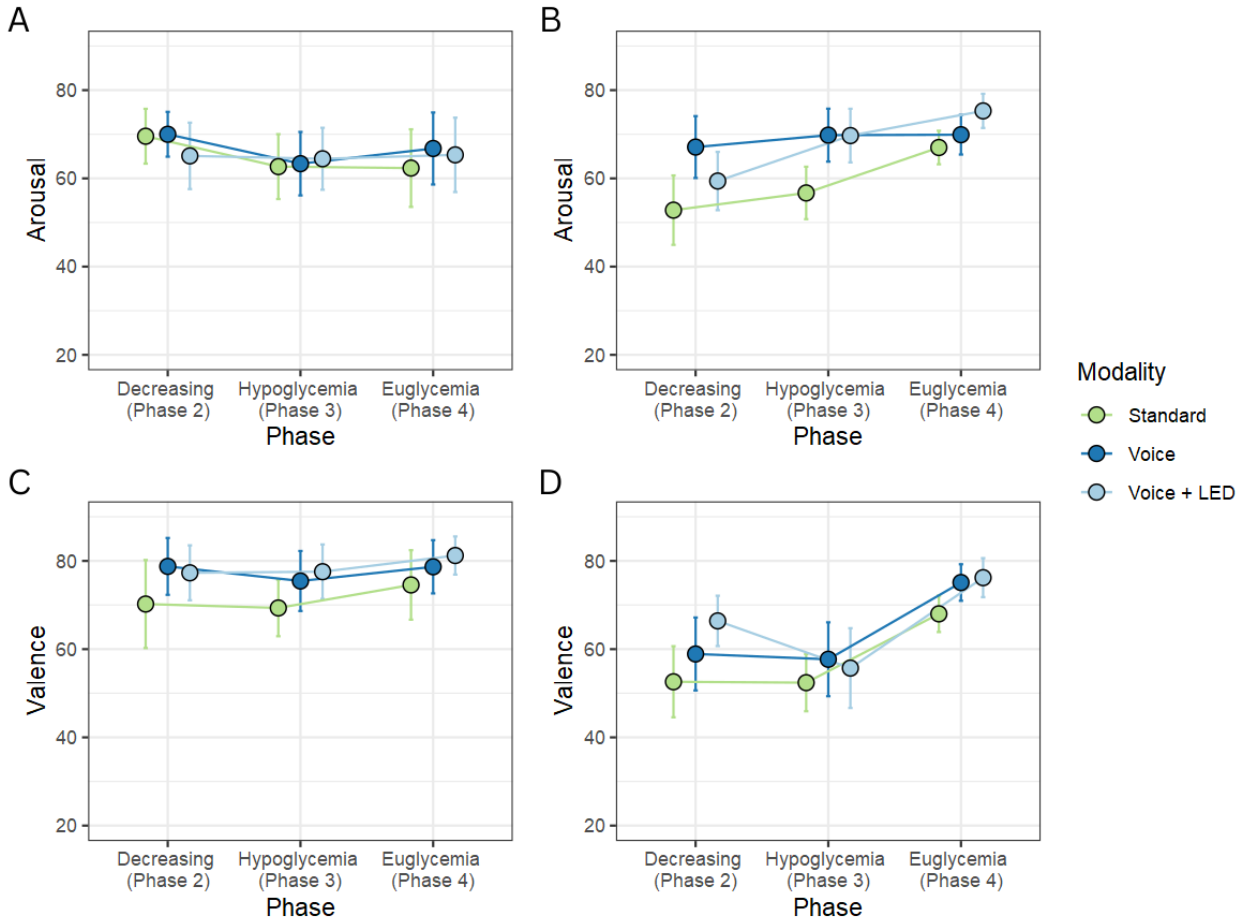
In this section, the results of the self-reported and physiological measures of emotional reaction (ie, self-reported valence and arousal and SCR) are described. A mixed-effect model was run on all these measures, with warning modality and blood glucose phase as independent variables.

Self-Reported Arousal and Valence

According to our results, the mixed-effect models were significant for both valence and arousal in both studies ($P<.001$). Figure 4 shows the means and SEs for arousal and valence.

XSL•FO
RenderX

Figure 4. Line plots of arousal and valence across warning modality and blood glucose phases across (A and C) simulated and (B and D) real-world driving. Error bars represent SEs.



In the simulated driving study, arousal was not significantly affected by either of the independent variables (modality: $F_{2,68}$=0.1; P=.91; partial η^2=0, blood glucose phase: $F_{2,68}$=1.1; P=.35; partial η^2=0.03). Valence was significantly affected by warning modality ($F_{2,68}$=3.9; P=.03; partial η^2=0.10) but not by blood glucose phase ($F_{2,68}$=1.1; P=.35; partial η^2=0.03).

In the real-world driving study, arousal was significantly affected both by warning modality ($F_{2,76}$=4.3; P=.02; partial η^2=0.10) and blood glucose phase ($F_{2,76}$=4.1; P=.02; partial η^2=0.10). Valence was significantly affected by blood glucose phase ($F_{2,76}$=9.3; P<.001; partial η^2=0.20) but not by warning modality ($F_{2,76}$=2; P=.14; partial η^2=0.05).

Physiological Arousal

According to our results, the mixed-effect models were not significant in both studies. Hence, the warning modality and blood glucose phase did not significantly affect physiological arousal measured via SCR neither in the simulated driving study (modality: $F_{2,68}$=2.46; P=.09; partial η^2=0.1, blood glucose phase: $F_{2,68}$=0.3; P=.74; partial η^2=0). nor in the real-world driving study (modality: $F_{2,76}$=0.8; P=.47; partial η^2=0, blood glucose phase: $F_{2,76}$=0.7; P=.5; partial η^2=0).

Technology Acceptance

In this section, the technology acceptance results (ie, Baldwin and Moore scales of urgency, effectiveness, and annoyance and the Van Der Laan Acceptance Scale) are described. Details are available in Table 3 [44].

XSL•FO
RenderX

Table 3. Technology acceptance measure across studies.

Measure and warning modalities	Simulated driving, mean (SD)	Real-world driving, mean (SD)
Urgency		
Standard	1.88 (1.55)	2.6 (1.26)
Voice	2.88 (1)	3 (0.82)
Voice+LED	3.12 (0.64)	3.6 (0.52)
Effectiveness		
Standard	2.25 (1.16)	3.3 (1.06)
Voice	2.75 (1.28)	3.6 (0.7)
Voice+LED	3.38 (1.06)	3.5 (0.71)
Annoyance		
Standard	2.25 (1.16)	1.7 (0.95)
Voice	0.88 (0.84)	1 (0.82)
Voice+LED	0.5 (0.76)	1.1 (1.2)
Acceptance		
Standard	3.35 (0.55)	3.52 (0.53)
Voice	3.74 (0.47)	3.88 (0.41)
Voice+LED	3.86 (0.47)	3.77 (0.61)

In the simulated driving study, the voice+LED modality elicited the highest sense of urgency and effectiveness, the least annoyance, and the highest acceptance, followed by the voice modality. In real-world driving, the voice+LED modality elicited the highest sense of urgency and least annoyance, while the voice modality elicited the most sense of effectiveness and the highest acceptance.

Preference Ranking

The average rank in the simulated driving study was 1.5 (SD 0.82) for the voice+LED modality, 2.2 (SD 0.6) for the voice modality, and 2.4 (SD 0.81) for the standard modality. The average rank in real-world driving was 1.8 (SD 0.79) for both the voice+LED and voice modalities and 2.4 (SD 0.84) for the standard modality.

In the real-world driving study, topical feedback showed that 6 participants mentioned that the light was not noticeable while driving (eg, “I have not noticed the light but at night, it certainly works better” [Participant 4]).

Discussion

Principal Findings

This study investigated the effect of warning modality (visual, vocal, and vocal with ambient lighting) on the emotional reaction and the acceptance of such a technology. Our results showed that voice warnings are more appreciated and considered more effective than standard warnings. However, the ambient lighting did affect such judgments.

Effects of Warning Modality on Emotional Reaction

Effect on SCR

No significant effect of warning modality (or blood glucose phase) on skin conductance was found. SCR measured through Empatica E4 has been previously shown to be linked with response to stimuli [45-47]. Moreover, it has been associated with blood glucose variation [48,49]. Therefore, we may consider the possibility that the measurement protocol used in this research may have experienced certain weaknesses and have affected the validity of the obtained results. Hence, we cannot consider the lack of significant results as negative evidence. Nevertheless, future research should further investigate the startling effect of the voice warning while driving, either by replicating our experimental setting, by using alternative electrodermal activity measurement tools [50], or by using other measures of emotional reaction, such as eye blinks [44].

Self-Reports

Our results showed that although in simulated driving the effect of modality on self-reported arousal was not significant, this was the case in real-world driving. In particular, higher arousal was observed during decreasing glucose and hypoglycemia for the voice and voice+LED modalities. Moreover, in the simulated driving study, the results on self-reported valence show a significant effect of modality, with voice+LED and voice warnings eliciting higher valence than a standard warning, particularly during decreasing glucose and hypoglycemia. This was not the case in the real-world driving study.

Despite the mixed results, the warning modality had a significant effect in the critical moments, that is, when the participants were about to experience or already experiencing hypoglycemia. Thus, our results showed the relevance of measuring emotional

XSL•FO
RenderX

reaction at different levels of blood glucose. While the Affective Slider is a very efficient measurement tool, it is important to note that some participants expressed a lack of confidence in self-reporting their emotions with it. Thus, future research might benefit from using alternative self-reported measures of emotion, such as the Positive and Negative Affect Schedule [51] or the Discreet Emotions Questionnaire [52].

Mixed Results

These mixed results preclude the formulation of definitive conclusions regarding the effect of modality on emotional reaction based on this study. As a sudden auditory stimulus can create a startling reaction, which can interfere with driving performance [10,33], future research should consider our recommendations and further investigate the design of a warning that is both effective and nonstartling.

Effects of Warning Modality on Acceptance and Preference

Acceptance

Our results demonstrated that the voice+LED modality tended to be the most valued regarding acceptance and preference. However, in the real-world driving study, this advantage, compared to the voice modality, seemed to decrease compared to the simulated driving study. This change might be due to the setup, where the ambient lighting was more visible in the laboratory than outdoors. Therefore, the advantage of the ambient lighting might have decreased in the real-world driving study. Thus, from the results, it is clear that the voice warning had an advantage over the standard warning, while the addition of ambient lighting (ie, voice+LED modality) did not bring a substantial advantage.

Preference

Finally, when asking participants for a posttest ranking of the warning modalities, results showed that voice (ie, both voice and voice+LED modalities) had a constant advantage over a beep with text (ie, standard modality). However, in the simulated driving study, the voice+LED modality was ranked first more often than in the real-world driving study, leading to interpret these results similar to the technology acceptance results, that is, adding ambient lighting to a voice warning was not considered substantially advantageous. These results might be influenced by the perception of the lights in daylight conditions, which differed between the simulated and the real-world driving settings. Based on the topical feedback from the participants experiencing the real-world driving setting, it might be that the contrast between the exterior and interior luminance was too low. As we did not measure the contrast in luminance between the daylight and the LEDs, future research should further investigate the use of ambient lighting as a component of in-vehicle warnings with greater control on luminance [53].

Implications

This paper involves implications both from health and automotive perceptive. First, our investigation represents a step forward in managing a real-life hazard associated with hypoglycemia. Our previous work [8] focused on designing an effective voice warning. This work compared it to a standard warning with an unspecified auditory signal with a text, and with an addition of ambient lighting, to make the voice warning less abrupt. Our results show an advantage for a voice warning (ie, spoken) over a tone with text but not for ambient lighting. While participants showed a higher preference for adding ambient lighting in the simulated driving study, the way the ambient lighting was set in the real-world driving study was not noticeable enough to replicate the results. Other work has investigated more attention-grabbing ambient lighting, such as blinking lights [54]. Future research should investigate if different typologies of ambient light patterns could affect emotional reaction and acceptance.

Second, our investigation aims to inspire in-vehicle technology designers to develop in-vehicle health warnings using the in-vehicle VA. While there are driver-state warnings, they predominantly alarm the driver with an unspecific beep and a text on the cockpit. Future research should investigate how using the in-vehicle VA to warn the driver about dizziness or lack of attention would be accepted by drivers. Moreover, providing health-related warnings while driving fits the concept of “health-conscious” cars [55,56]. Along the lines of our investigation, there have been some attempts to develop blood glucose monitoring interfaces for the car [17,18]. However, they primarily rely on visual displays and are ill-adapted to the context of driving.

Finally, this study assumed the delivery of a hypoglycemia warning in a car with an autonomy of level 0 (ie, no automation) or level 1 (ie, with driver assistance). As cars are becoming increasingly automated, a hypoglycemia warning should be compatible with cars with a higher level of autonomy. However, the warning designed in this work is compatible with higher levels of automation. For instance, during autonomous driving, the in-vehicle VA could alert the driver that hypoglycemia has been detected and trigger the car to autonomously stop. During manual driving, the in-vehicle VA could warn the driver and trigger the car to take over (switch from manual to autonomous driving) and pull over.

Limitations and Future Research

Despite our best efforts, this investigation involved certain limitations. First, the sample size was rather small. Nevertheless, valuable insights on digital solutions can be provided with a small sample size [57-59]. Thus, although it does not allow drawing conclusions on the interaction of drivers with T1DM with in-vehicle hypoglycemia warnings, it motivates further research in this domain.

Second, emotional response to the warning was evaluated in different blood glucose states and a controlled setting (ie, simulator and closed circuit). However, the warning was not delivered when relevant (ie, only when the driver was actually undergoing hypoglycemia) or while the participant was driving on a public road unaware of the upcoming critical state. While our method allowed controlling for the blood glucose in the emotional response, it did not allow us to find the opportune moment for intervention delivery, that is, at what point of upcoming hypoglycemia is the warning most appropriate (both in terms of emotional reaction and acceptance). Future research should investigate the effectiveness of such an intervention in

XSL•FO
RenderX

a more realistic context, where the driver does not expect to be warned and is actually about to undergo hypoglycemia while on a public road.

Conclusions

This paper proposes the use of the in-vehicle VA and ambient lighting system to deliver a hypoglycemia warning, ensuring a hands-free alert. The investigation focused on the extent to which warning modality could affect emotional response and acceptance both in a simulated and real-world environment. Although further investigations are needed, our results suggest, together with our previous work [8], that implementing multimodal warnings can improve the management of hypoglycemia in cars and also emphasize the potential of in-vehicle VA for delivering health-related warnings.

Acknowledgments

This work was part-funded by the Swiss National Science Foundation, Sinergia-Project CRSII5_183569 “Design and Evaluation of a Vehicle Hypoglycaemia Warning System in Diabetes (HEADWIND Project).”

Data Availability

The data sets generated and analyzed during this study are available from the corresponding author on reasonable request.

Authors' Contributions

CB, ABK, and TK were responsible for the warning design and development. CB, MM, FW, and TK were responsible for the design of the driving scenarios. CB, MM, and VFL were responsible for the data collection on the intervention. CB was responsible for the data analysis and the first draft of this paper. All authors were responsible for critical feedback and final revisions of the paper.

Conflicts of Interest

TK is affiliated with the Centre for Digital Health Interventions, a joint initiative of the Institute for Implementation Science in Health Care, University of Zurich; the Department of Management, Technology, and Economics at ETH Zurich; and the Institute of Technology Management and School of Medicine at the University of St.Gallen. The Centre for Digital Health Interventions is funded in part by CSS, a Swiss health insurer; Mavie Next, an Austrian health insurer; and MTIP, a Swiss digital health investor. TK is also a cofounder of Pathmate Technologies, a university spin-off company that creates and delivers digital clinical pathways. However, neither CSS, Mavie Next, MTIP nor Pathmate Technologies were involved in this research.

References

1. Green A, Hede SM, Patterson CC, Wild SH, Imperatore G, Roglic G, et al. Type 1 diabetes in 2017: global estimates of incident and prevalent cases in children and adults. Diabetologia 2021;64(12):2741-2750 [FREE Full text] [doi: 10.1007/s00125-021-05571-8] [Medline: 34599655]
2. Allen KV, Pickering MJ, Zammitt NN, Hartsuiker RJ, Traxler MJ, Frier BM, et al. Effects of acute hypoglycemia on working memory and language processing in adults with and without type 1 diabetes. Diabetes Care 2015;38(6):1108-1115 [FREE Full text] [doi: 10.2337/dc14-1657] [Medline: 25758768]
3. Warren RE, Frier BM. Hypoglycaemia and cognitive function. Diabetes Obes Metab 2005;7(5):493-503. [doi: 10.1111/j.1463-1326.2004.00421.x] [Medline: 16050942]
4. Cryer PE. Hypoglycemia: Pathophysiology, Diagnosis, and Treatment. Oxford: Oxford University Press; 1997.
5. Blumer I, Kenshole AB, Stilman J, Lewis GF. Insulin-treated diabetes and driving: legal jeopardy and consequences of hypoglycemia. Can J Diabetes 2019;43(3):221-223. [doi: 10.1016/j.jcjd.2018.08.201] [Medline: 30348593]
6. Cox DJ, Gonder-Frederick LA, Kovatchev BP, Clarke WL. Self-treatment of hypoglycemia while driving. Diabetes Res Clin Pract 2001;54(1):17-26. [doi: 10.1016/s0168-8227(01)00274-1] [Medline: 11532326]
7. Cox DJ, Penberthy JK, Zrebiec J, Weinger K, Aikens JE, Frier B, et al. Diabetes and driving mishaps: frequency and correlations from a multinational survey. Diabetes Care 2003;26(8):2329-2334 [FREE Full text] [doi: 10.2337/diacare.26.8.2329] [Medline: 12882857]
8. Bérubé C, Lehmann VF, Maritsch M, Kraus M, Feuerriegel S, Wortmann F, et al. Effectiveness and user perception of an in-vehicle voice warning for hypoglycemia: development and feasibility trial. JMIR Hum Factors 2024 Jan 09;11:e42823 [FREE Full text] [doi: 10.2196/42823] [Medline: 38194257]
9. Landry S, Jeon M, Lautala P, Nelson D. Design and assessment of in-vehicle auditory alerts for highway-rail grade crossings. Transp Res F: Traffic Psychol Behav 2019;62:228-245. [doi: 10.1016/j.trf.2018.12.024]
10. Biondi F, Rossi R, Gastaldi M, Mulatti C. Beeping ADAS: reflexive effect on drivers' behavior. Transp Res F: Traffic Psychol Behav 2014;25(A):27-33 [FREE Full text] [doi: 10.1016/j.trf.2014.04.020]

XSL•FO
RenderX

11. Larsson P, Opperud A, Fredriksson K, Västfjäll D. Emotional and behavioural response to auditory icons and earcons in driver-vehicle interfaces. 2009 Presented at: Proceedings of the 21st International Technical Conference on the Enhanced Safety of Vehicles; June 2009; Stuttgart, Germany URL: https://tinyurl.com/msz254ve
12. Larsson P, Västfjäll D. Emotional and behavioural responses to auditory interfaces in commercial vehicles. IJVNV 2013;9(1-2):75-95. [doi: 10.1504/ijvnv.2013.053818]
13. Caberletti L, Elfmann K, Kummel M, Schierz C. Influence of ambient lighting in a vehicle interior on the driver's perceptions. Light Res Technol 2010;42(3):297-311. [doi: 10.1177/1477153510370554]
14. Kim T, Kim Y, Jeon H, Choi CS, Suk HJC. Emotional response to in-car dynamic lighting. Int J Automot Technol 2021;22(4):1035-1043. [doi: 10.1007/s12239-021-0093-4]
15. Löcken A, Heuten W, Boll S. Enlightening drivers: a survey on in-vehicle light displays. 2016 Presented at: Automotive'UI 16: Proceedings of the 8th International Conference on Automotive User Interfaces and Interactive Vehicular Applications; October 24-26, 2016; Ann Arbor, MI, USA p. 97-104 URL: https://dl.acm.org/doi/proceedings/10.1145/3003715 [doi: 10.1145/3003715.3005416]
16. McCartt AT, Hellinga LA, Bratiman KA. Cell phones and driving: review of research. Traffic Inj Prev 2006;7(2):89-106. [doi: 10.1080/15389580600651103] [Medline: 16854702]
17. Kerr D, Olateju T. Driving with diabetes in the future: in-vehicle medical monitoring. J Diabetes Sci Technol 2010;4(2):464-469 [FREE Full text] [doi: 10.1177/193229681000400228] [Medline: 20307408]
18. Hoskins M. Diabetes data behind the wheel... #WeAreNotWaiting. Healthline Media. 2018. URL: https://web.archive.org/web/20201112015408/https://www.healthline.com/diabetesmine/diabetes-data-in-cars-wearenotwaiting#1 [accessed 2024-05-04]
19. Advanced driver assistance systems. European Commission. 2016. URL: https://web.archive.org/web/20231225024440/https://road-safety.transport.ec.europa.eu/system/files/2021-07/ersosynthesis2016-adas15_en.pdf [accessed 2024-04-05]
20. Baldwin CL, Moore C. Perceived urgency, alerting effectiveness and annoyance of verbal collision avoidance system messages. Proc Hum Factors Ergon Soc Annu Meet 2016;46(22):1848-1852. [doi: 10.1177/154193120204602215]
21. Statement of principles on the design of high-priority warning signals for in-vehicle intelligent transport systems—draft. International Harmonized Research Activities (IHRA) Working Group on Intelligent Transport Systems (ITS). 2009. URL: https://unece.org/DAM/trans/doc/2009/wp29/ITS-17-03e.pdf [accessed 2024-05-04]
22. Kinsella B, Mutchler A. In-car voice assistant consumer adoption report 2019. Voicebot.ai. 2019. URL: https://web.archive.org/web/20231002110008/https://voicebot.ai/in-car-voice-assistant-consumer-adoption-report-2019/ [accessed 2024-04-05]
23. Kinsella B, Mutchler A. In-car voice assistant consumer adoption report 2020. Voicebot.ai. 2020. URL: https://web.archive.org/web/20231002082700/https://voicebot.ai/in-car-voice-assistant-consumer-adoption-report-2020/ [accessed 2024-04-05]
24. Kinsella B. In-car voice assistant users 127 million in the U.S. with strong adoption in the UK and Germany as well—new report. Voicebot.ai. 2021. URL: https://web.archive.org/web/20230221102553/https://voicebot.ai/2021/08/23/in-car-voice-assistant-users-127-million-in-the-u-s-with-strong-adoption-in-the-uk-and-germany-as-well-new-report/ [accessed 2021-08-23]
25. Strayer DL, Cooper JM, Goethe RM, McCarty MM, Getty DJ, Biondi F. Assessing the visual and cognitive demands of in-vehicle information systems. Cogn Res Princ Implic 2019;4(1):18 [FREE Full text] [doi: 10.1186/s41235-019-0166-3] [Medline: 31227955]
26. Nass C, Steuer J, Tauber ER. Computers are social actors. 1994 Presented at: CHI '94: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems; April 24-28, 1994; Boston, MA, USA p. 72-78 URL: https://dl.acm.org/doi/proceedings/10.1145/191666 [doi: 10.1145/191666.191703]
27. Cho E, Molina MD, Wang J. The effects of modality, device, and task differences on perceived human likeness of voice-activated virtual assistants. Cyberpsychol Behav Soc Netw 2019;22(8):515-520. [doi: 10.1089/cyber.2018.0571] [Medline: 31314579]
28. Mahajan K, Large DR, Burnett G, Velaga NR. Exploring the effectiveness of a digital voice assistant to maintain driver alertness in partially automated vehicles. Traffic Inj Prev 2021;22(5):378-383. [doi: 10.1080/15389588.2021.1904138] [Medline: 33881365]
29. Voice-powered assistants make way for integrated, intuitive co-pilots. Cerence. 2022. URL: https://tinyurl.com/45v8cfau [accessed 2022-01-12]
30. Hofmann H, Hermanutz M, Tobisch V, Ehrlich U, Berton A, Minker W. Evaluation of in-car SDS notification concepts for incoming proactive events. In: Situated Dialog in Speech-Based Human-Computer Interaction. Cham: Springer; 2016:111-124.
31. Schmidt M, Minker W, Werner S. How users react to proactive voice assistant behavior while driving. 2020 Presented at: Proceedings of the Twelfth Language Resources and Evaluation Conference; May 2020; Marseille, France p. 485-490 URL: https://api.semanticscholar.org/CorpusID:218973946
32. Ringfort-Felner R, Laschke M, Sadeghian S, Hassenzahl M. Kiro: a design fiction to explore social conversation with voice assistants. Proc ACM Hum-Comput Interact 2022;6(GROUP):1-21. [doi: 10.1145/3492852]

XSL•FO
RenderX

33. Woodhead MM. Effect of brief loud noise on decision making. J Acoust Soc Am 1959;31(10):1329-1331. [doi: 10.1121/1.1907629]
34. Heise T, Koschinsky T, Heinemann L, Lodwig V, Glucose Monitoring Study Group. Hypoglycemia warning signal and glucose sensors: requirements and concepts. Diabetes Technol Ther 2003;5(4):563-571. [doi: 10.1089/152091503322250587] [Medline: 14511411]
35. Löcken A, Heuten W, Boll SC. AutoAmbiCar: using ambient light to inform drivers about intentions of their automated cars. 2016 Presented at: AutomotiveUI '16 Adjunct: Adjunct Proceedings of the 8th International Conference on Automotive User Interfaces and Interactive Vehicular Applications; October 24-26, 2016; Ann Arbor, MI, USA p. 57-62 URL: https://dl.acm.org/doi/proceedings/10.1145/3004323 [doi: 10.1145/3004323.3004329]
36. Parasuraman A, Colby CL. An updated and streamlined Technology Readiness Index: TRI 2.0. J Serv Res 2014;18(1):59-74. [doi: 10.1177/1094670514539730]
37. Lehmann V, Zueger T, Maritsch M, Kraus M, Albrecht C, Bérubé C, et al. Machine learning for non-invasive sensing of hypoglycaemia while driving in people with diabetes. Diabetes Obes Metab 2023;25(6):1668-1676 [FREE Full text] [doi: 10.1111/dom.15021] [Medline: 36789962]
38. Bérubé C. GitHub—hypogycemia-warning. catbot0. 2021. URL: https://github.com/catbot0/hypogycemia-warning [accessed 2024-03-15]
39. Betella A, Verschure PFMJ. The Affective Slider: a digital self-assessment scale for the measurement of human emotions. PLoS One 2016;11(2):e0148037 [FREE Full text] [doi: 10.1371/journal.pone.0148037] [Medline: 26849361]
40. Van Der Laan JD, Heino A, De Waard D. A simple procedure for the assessment of acceptance of advanced transport telematics. Transp Res Part C Emerg Technol 1997;5(1):1-10. [doi: 10.1016/s0968-090x(96)00025-3]
41. Dahlbäck N, Jönsson A, Ahrenberg L. Wizard of Oz studies—why and how. Knowl-Based Syst 1993;6(4):258-266. [doi: 10.1016/0950-7051(93)90017-n]
42. Boucsein W. Electrodermal Activity. 2nd Edition. New York: Springer; 2012.
43. Braithwaite JJ, Watson DG, Jones R, Rowe ML. A guide for analysing electrodermal activity (EDA) and skin conductance responses (SCRs) for psychological experiments. Psychophysiology. 2013. URL: https://www.lancaster.ac.uk/media/lancaster-university/content-assets/documents/psychology/ABriefGuideforAnalysingElectrodermalActivityv4.pdf) [accessed 2024-03-15]
44. Brown P, Rothwell JC, Thompson PD, Britton TC, Day BL, Marsden CD. New observations on the normal auditory startle reflex in man. Brain 1991;114((Pt 4)):1891-1902. [doi: 10.1093/brain/114.4.1891] [Medline: 1884184]
45. Borrego A, Latorre J, Alcañiz M, Llorens R. Reliability of the Empatica E4 wristband to measure electrodermal activity to emotional stimuli. 2019 Presented at: 2019 International Conference on Virtual Rehabilitation (ICVR); July 21-24, 2019; Tel Aviv, Israel p. 1-2 URL: https://ieeexplore.ieee.org/document/8994546/ [doi: 10.1109/icvr46560.2019.8994546]
46. Milstein N, Gordon I. Validating measures of electrodermal activity and heart rate variability derived from the Empatica E4 utilized in research settings that involve interactive dyadic states. Front Behav Neurosci 2020;14:148 [FREE Full text] [doi: 10.3389/fnbeh.2020.00148] [Medline: 33013337]
47. Ragot M, Martin N, Em S, Pallamin N, Diverrez JM. Emotion recognition using physiological signals: laboratory vs. wearable sensors. 2017 Presented at: Advances in Human Factors in Wearable Technologies and Game Design: Proceedings of the AHFE 2017 International Conference on Advances in Human Factors and Wearable Technologies; July 17-21, 2017; The Westin Bonaventure Hotel, Los Angeles, CA, USA p. 15-220 URL: https://doi.org/10.1007/978-3-319-60639-2_2 [doi: 10.1007/978-3-319-60639-2_2]
48. Bent B, Cho PJ, Henriquez M, Wittmann A, Thacker C, Feinglos M, et al. Engineering digital biomarkers of interstitial glucose from noninvasive smartwatches. NPJ Digit Med 2021;4(1):89 [FREE Full text] [doi: 10.1038/s41746-021-00465-w] [Medline: 34079049]
49. Bogue-Jimenez B, Huang X, Powell D, Doblas A. Selection of noninvasive features in wrist-based wearable sensors to predict blood glucose concentrations using machine learning algorithms. Sensors (Basel) 2022;22(9):3534 [FREE Full text] [doi: 10.3390/s22093534] [Medline: 35591223]
50. Kappeler-Setz C, Gravenhorst F, Schumm J, Arnrich B, Tröster G. Towards long term monitoring of electrodermal activity in daily life. Pers Ubiquitous Comput 2013;17(2):261-271 [FREE Full text] [doi: 10.1007/s00779-011-0463-4]
51. Crawford JR, Henry JD. The Positive and Negative Affect Schedule (PANAS): construct validity, measurement properties and normative data in a large non-clinical sample. Br J Clin Psychol 2004;43(Pt 3):245-265. [doi: 10.1348/0144665031752934] [Medline: 15333231]
52. Harmon-Jones C, Bastian B, Harmon-Jones E. The Discrete Emotions Questionnaire: a new tool for measuring state self-reported emotions. PLoS One 2016;11(8):e0159915 [FREE Full text] [doi: 10.1371/journal.pone.0159915] [Medline: 27500829]
53. Blankenbach K, Hertlein F, Hoffmann S. Advances in automotive interior lighting concerning new LED approach and optical performance. J Soc Info Display 2020;28(8):655-667 [FREE Full text] [doi: 10.1002/jsid.887]
54. Kim T, Lee G, Park M, Lee HM, Park JW, Suk HJ. User responses to dynamic light in automobiles with EEG and self-assessments. IEEE Access 2022;10:123847-123857 [FREE Full text] [doi: 10.1109/access.2022.3223665]

XSL•FO
RenderX

55. Koch K, Mishra V, Liu S, Berger T, Fleisch E, Kotz D, et al. When do drivers interact with in-vehicle well-being interventions? An exploratory analysis of a longitudinal study on public roads. Proc ACM Interact Mob Wearable Ubiquitous Technol 2021;5(1):19 [FREE Full text] [doi: 10.1145/3448116] [Medline: 35178497]
56. Koch K, Tiefenbeck V, Liu S, Berger T, Fleisch E, Wortmann F. Taking mental health and well-being to the streets: an exploratory evaluation of in-vehicle interventions in the wild. 2021 Presented at: CHI '21: Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems; May 8-13, 2021; Yokohama, Japan p. 1-15 URL: https://dl.acm.org/doi/10.1145/3411764.3446865 [doi: 10.1145/3411764.3446865]
57. Papathanail I, Vasiloglou MF, Stathopoulou T, Ghosh A, Baumann M, Faeh D, et al. A feasibility study to assess Mediterranean diet adherence using an AI-powered system. Sci Rep 2022;12(1):17008 [FREE Full text] [doi: 10.1038/s41598-022-21421-y] [Medline: 36220998]
58. Payne O, Smith V, Rolnik DL, Davies-Tuck M, Warty R, Seo D, et al. Virtual reality and its use in post-operative pain following laparoscopy: a feasibility study. Sci Rep 2022;12(1):13137 [FREE Full text] [doi: 10.1038/s41598-022-17183-2] [Medline: 35907935]
59. Heijmans M, Habets JGV, Herff C, Aarts J, Stevens A, Kuijf ML, et al. Monitoring Parkinson's disease symptoms during daily life: a feasibility study. NPJ Parkinsons Dis 2019;5:21 [FREE Full text] [doi: 10.1038/s41531-019-0093-5] [Medline: 31583270]

Abbreviations

SCR: skin conductance response
T1DM: type 1 diabetes mellitus
VA: voice assistant

Edited by A Kushniruk; submitted 08.03.23; peer-reviewed by K Baker, Y Jeem; comments to author 03.05.23; revised version received 23.05.23; accepted 02.03.24; published 18.04.24.

Please cite as:
Bérubé C, Maritsch M, Lehmann VF, Kraus M, Feuerriegel S, Züger T, Wortmann F, Stettler C, Fleisch E, Kocaballi AB, Kowatsch T
Multimodal In-Vehicle Hypoglycemia Warning for Drivers With Type 1 Diabetes: Design and Evaluation in Simulated and Real-World Driving
JMIR Hum Factors 2024;11:e46967
URL: https://humanfactors.jmir.org/2024/1/e46967
doi:10.2196/46967
PMID:38635313

©Caterina Bérubé, Martin Maritsch, Vera Franziska Lehmann, Mathias Kraus, Stefan Feuerriegel, Thomas Züger, Felix Wortmann, Christoph Stettler, Elgar Fleisch, A Baki Kocaballi, Tobias Kowatsch. Originally published in JMIR Human Factors (https://humanfactors.jmir.org), 18.04.2024. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on https://humanfactors.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Usability Comparison Among Healthy Participants of an Anthropomorphic Digital Human and a Text-Based Chatbot as a Responder to Questions on Mental Health: Randomized Controlled Trial

Almira Osmanovic Thunström[1,2], MSc; Hanne Krage Carlsen[2,3], MPH, PhD; Lilas Ali[1,4,5], RN, PhD; Tomas Larson[1,2], MSc, PhD; Andreas Hellström[6], PhD; Steinn Steingrimsson[1,2], BSc, MD, PhD

[1]Region Västra Götaland, Psychiatric Department, Sahlgrenska University Hospital, Gothenburg, Sweden

[2]Section of Psychiatry and Neurochemistry, Institute of Neuroscience and Physiology, Sahlgrenska Academy, University of Gothenburg, Gothenburg, Sweden

[3]Region Västra Götaland, Centre of Registers, Gothenburg, Sweden

[4]Institute of Health Care Sciences, Centre for Person-Centred Care, Sahlgrenska Academy, University of Gothenburg, Gothenburg, Sweden

[5]Centre for Person-Centred Care, University of Gothenburg, Gothenburg, Sweden

[6]Department of Technology Management and Economics, Chalmers University of Technology, Gothenburg, Sweden

Corresponding Author:
Steinn Steingrimsson, BSc, MD, PhD
Region Västra Götaland
Psychiatric Department
Sahlgrenska University Hospital
Journalvägen 5
Gothenburg, 41650
Sweden
Phone: 46 313421000
Email: steinn.steingrimsson@gu.se

Abstract

Background: The use of chatbots in mental health support has increased exponentially in recent years, with studies showing that they may be effective in treating mental health problems. More recently, the use of visual avatars called digital humans has been introduced. Digital humans have the capability to use facial expressions as another dimension in human-computer interactions. It is important to study the difference in emotional response and usability preferences between text-based chatbots and digital humans for interacting with mental health services.

Objective: This study aims to explore to what extent a digital human interface and a text-only chatbot interface differed in usability when tested by healthy participants, using BETSY (Behavior, Emotion, Therapy System, and You) which uses 2 distinct interfaces: a digital human with anthropomorphic features and a text-only user interface. We also set out to explore how chatbot-generated conversations on mental health (specific to each interface) affected self-reported feelings and biometrics.

Methods: We explored to what extent a digital human with anthropomorphic features differed from a traditional text-only chatbot regarding perception of usability through the System Usability Scale, emotional reactions through electroencephalography, and feelings of closeness. Healthy participants (n=45) were randomized to 2 groups that used a digital human with anthropomorphic features (n=25) or a text-only chatbot with no such features (n=20). The groups were compared by linear regression analysis and *t* tests.

Results: No differences were observed between the text-only and digital human groups regarding demographic features. The mean System Usability Scale score was 75.34 (SD 10.01; range 57-90) for the text-only chatbot versus 64.80 (SD 14.14; range 40-90) for the digital human interface. Both groups scored their respective chatbot interfaces as average or above average in usability. Women were more likely to report feeling annoyed by BETSY.

Conclusions: The text-only chatbot was perceived as significantly more user-friendly than the digital human, although there were no significant differences in electroencephalography measurements. Male participants exhibited lower levels of annoyance with both interfaces, contrary to previously reported findings.

XSL•FO
RenderX

(JMIR Hum Factors 2024;11:e54581) doi:10.2196/54581

KEYWORDS

chatbot; chatbots; chat-bot; chat-bots; text-only chatbot, voice-only chatbot; mental health; mental illness; mental disease; mental diseases; mental illnesses; mental health service; mental health services; interface; system usability; usability; digital health; machine learning; ML; artificial intelligence; AI; algorithm; algorithms; NLP; natural language processing

Introduction

Conversational user interfaces, also known as chatbots, have been a part of human-computer interactions since the 1960s. Most notable and one of the earliest examples is the ELIZA system, which aimed at simulating a human psychologist [1,2]. A subsequent system named PARRY was implemented in 1972, in which the conversational agent was designed to emulate a patient experiencing schizophrenia [3]. It is by no means a coincidence that the 2 earliest systems targeting the replication of human behavior through natural language processing were both derived from the field of psychiatry. The use of conversational agents has increased exponentially in the past decade [4]. With the availability of systems and the increasing need for 24-hour availability due to globalization, Radziwill and Benton [5] found that perhaps as many as 1 of 3 web-based conversations were conducted with a chatbot or a system moderated by language models, of which some have garnered more than 100 million users [4-8].

Previous research on rule-based conversational agents has shown promise with respect to the alleviation of mental health problems [4,9-11]. In a study by Oh et al [12], patients with panic disorder were randomized to support via a chatbot or support via a self-help book. The patients who were assigned to a chatbot as a support system for exercises in cognitive behavioral therapy were more likely to show symptom alleviation [12]. Digital evaluations as well as digital deliverance of mental health aid were more intensively explored following the COVID-19 pandemic [13]. In a study by Islam et al [14], the authors explored a similar design to that of Oh et al [12] and randomized a set of participants to either book or chatbot intervention for support regarding mental health issues [14]. The group of participants allocated to the chatbot intervention also significantly improved control of helplessness and social phobia scores. Some studies have shown that even a single exposure to a chatbot therapist can have a positive influence on the current state of well-being and repeated exposure can be a good complementary treatment for anxiety [9,10,15,16].

In recent years, a novel facet has been introduced into the evaluative framework for mental health chatbots: the incorporation of voice-controlled visual avatars embodying humanoid countenances colloquially referred to as “digital humans.” These digital entities harness the power of machine learning, emotion-infused linguistics, and adept emulation of facial expressions to cultivate a profound emotional rapport with their users. Research shows that human features elicit more social engagement and can trigger a stronger emotional bond [17]. This has primarily been measured through electroencephalography (EEG) with a specific focus being placed on the importance of increased α and θ wave activity as indicators of overall emotional stability and positive response to stimuli [18-20], while β wave activity has largely been associated with less desirable states of mind such as anxiety and an active stress response [19,21].

While acknowledging the inherent complexity of brain states and wave activity, delving into the extent to which distinct brain wave frequencies exert influence during a chat session presents an intriguing avenue for investigating the emotional states of the user. In a study by Bos et al [22], the authors explored capturing vigilance and states of emotion with EEG in usability testing of chatbot technology. The study findings revealed that EEG effectively captured the facets of user experience and conversation that piqued interest. This was accomplished through the delineation of γ wave activity, predominantly linked with positivity and problem-solving. Consequently, this approach affords researchers a more objective means of apprehending user experience. A study by Ciechanowski et al [23] indicated that there is a difference in emotional response and usability preference between text-based chatbots and digital humans, with text-based chatbots eliciting more positive interactions.

Although EEG has served as a proficient tool for quantifying objective assessments of emotional responses to chatbot interventions, it is customary to use usability scales for capturing the subjective dimensions pertaining to emotions and experiences in the context of chatbot systems. While a universally accepted benchmark for conducting usability tests on chatbots remains elusive, numerous studies have gravitated toward the adoption of the System Usability Scale (SUS-10) [24-29] and the Speech User Interface Service Quality scale. SUS-10 captures the overall usability of a system independently of the platform or interface. The score ranges from 0 to 100, indicating higher usability with increasing scores [26]. A score of 68 is considered as a passing grade, while a score below 50 is considered as indicating that the system has less optimal usability. For a system to be considered as exceptionally good in terms of its design and usability, a score of 85 on average should be applied [29-31]. In the previously mentioned study by Oh et al [12], mean SUS-10 was not significantly worse or better comparing a chatbot and a book: 64.5 (SD 17.0) versus 69.5 (SD 17.2), respectively (P=.35). Several studies have advocated the idea that chatbots represent user-friendly alternatives to conventional analog methods or standard digital tools, such as forms [4]. Nonetheless, there is research that suggests the design flaws in a chatbot system can markedly diminish its effectiveness, potentially leading to perceptions of unhelpfulness among users [3]. Chatbots that are perceived as unhelpful, repetitive, or lacking the users’ trust tend to receive a lower SUS-10 score [32].

Many social chatbots aim to comfort, support, and advise their users [3]. Studies show that the availability of chatbot technology is what is central to its perception of usefulness

XSL•FO
RenderX

compared with human therapists. However, studies have also noted that most users prefer human therapists and are more interested in using the system as a complementary tool when a human therapist is not available [33-35]. While mental health chatbots are generally viewed positively by the user, there are many issues that can lead to decreased usability, lower SUS-10 scores, and undesirable outcomes such as irritation or worsened mental health. The propensity for misunderstanding, miscommunication, and annoyance are frequently reported in qualitative assessments of social support chatbots [33-35]. Feeling annoyed by repetitive messaging, incoherent conversations, and inability to comprehend the user's needs are frequently named as issues that increase the feeling of annoyance in users of social support chatbots [34]. The selection of an interface can wield a considerable influence on both the effectiveness and user-friendliness of a system. Users exhibit disparate reactions to chatbots depending on whether they incorporate an avatar, particularly one with humanoid attributes capable of evoking emotions. Although our understanding of chatbot usability and user preference is somewhat limited, investigations into anthropomorphic interfaces do underscore their ability to affect our emotional states [36].

The chatbot used in our study, known as BETSY (Behavior, Emotion, Therapy System, and You), uses 2 distinct interfaces: a digital human, voice-activated user interface with anthropomorphic features and a text-only user interface. Within the scope of our investigation, we aimed to thoroughly examine both interface modalities. Phase 1 of usability testing involves enlisting the participation of healthy volunteers.

The aim of this study was to explore to what extent a digital human and a text-only chatbot interface differed in usability when tested by healthy participants. We also set out to explore how chatbot-generated conversations on mental health (specific to each interface) affected self-reported feelings and biometrics.

Methods

Construction of the User Interface (BETSY)

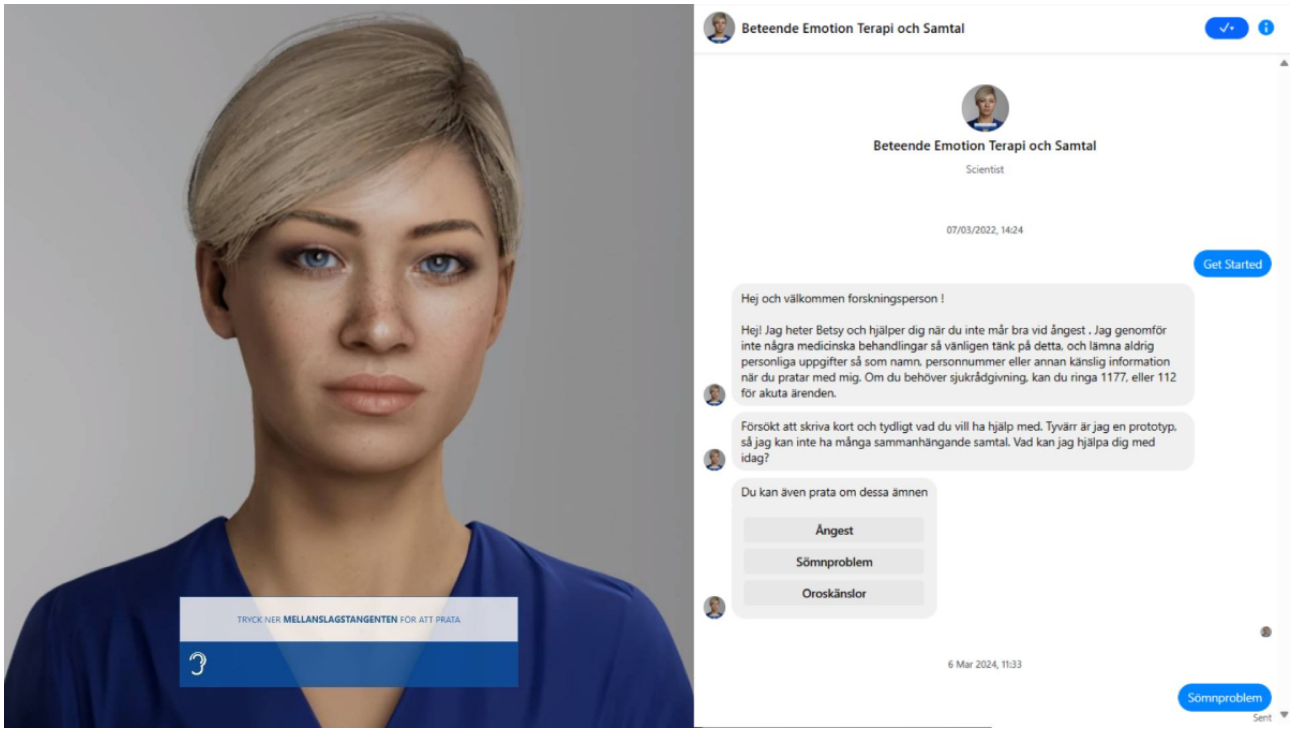
This project adopted a participatory design approach to ensure the broad involvement of health care professionals, patients, and the public. A multidisciplinary team consisting of 2 psychiatrists, 2 psychiatric nurses, 4 clinical psychologists, 1 user of health care services, and 1 engineer was assembled to comprehensively address ethical, medical, and legal considerations for a potential chatbot. Team members were selected for their expertise in digitalization and psychiatry. Before the initial workshop, where the algorithm's preliminary outline was presented, the engineer created a survey. This survey drew partly from Radziwill and Benton [5] quality attribute listing, which synthesized findings from various chatbot usability projects.

A survey was distributed to the public via the secure research platform Psytoolkit.org, offering heightened anonymity by omitting the collection of metadata such as IP addresses and locations. The survey comprised 8 multiple-choice questions and 4 open-ended free-text questions, covering demographic information, design requirements, functionality suggestions, and overall attitudes toward mental health chatbots. It was accessible for 14 days and disseminated through various social media channels. Subsequently, the collected data were analyzed to inform a series of 4 workshops conducted by the group between June 2020 and December 2020. During these workshops, the chatbot's design, encompassing appearance, content, and personality, underwent iterative development based on input from the general public and co-designer feedback, with the latter representing a patient perspective. A comprehensive account of this process will be available in a separate publication.

Two versions of the chatbot (Figure 1) were created: one enabling voice interaction with a facial expression and an avatar component, and another relying solely on text-based communication with an avatar image. The digital human was implemented using Dialogflow (Google) for conversation logic and connected to the UNEEQ platform for the human-avatar interface. Data infrastructure was hosted by Deloitte Digital and VästraGötalandsregionen/VGR-IT. In contrast, the text-only BETSY chatbot was developed on the Itsalive.io platform and deployed to a research and development account on Facebook that was closed to the public. Importantly, no personal metadata were collected during on-site testing via digital platforms. The users did not use their personal social media accounts to talk to the chatbot.

XSL•FO
RenderX

Figure 1. Two versions of BETSY (Behavior, Emotion, Therapy System, and You): digital human, voice-activated (left) and text-only (right).



Both versions of BETSY encompassed 24 topics (detailed in Multimedia Appendix 1) related to mental health, including anxiety, depression, stress, sleep, addiction, eating disorders, anger, hopelessness, helplessness, loneliness, sadness, suicidal ideation, and suicidality, among others. These chatbots were designed in the Swedish language. An assessment was conducted to evaluate the alignment of the text-only and digital human algorithms. Specifically, testers posed identical questions to both systems within various domains, with only 1 instance revealing a discrepancy when the digital human could not provide an appropriate response while the text-based bot could, indicating the need for further refinement.

Recruitment

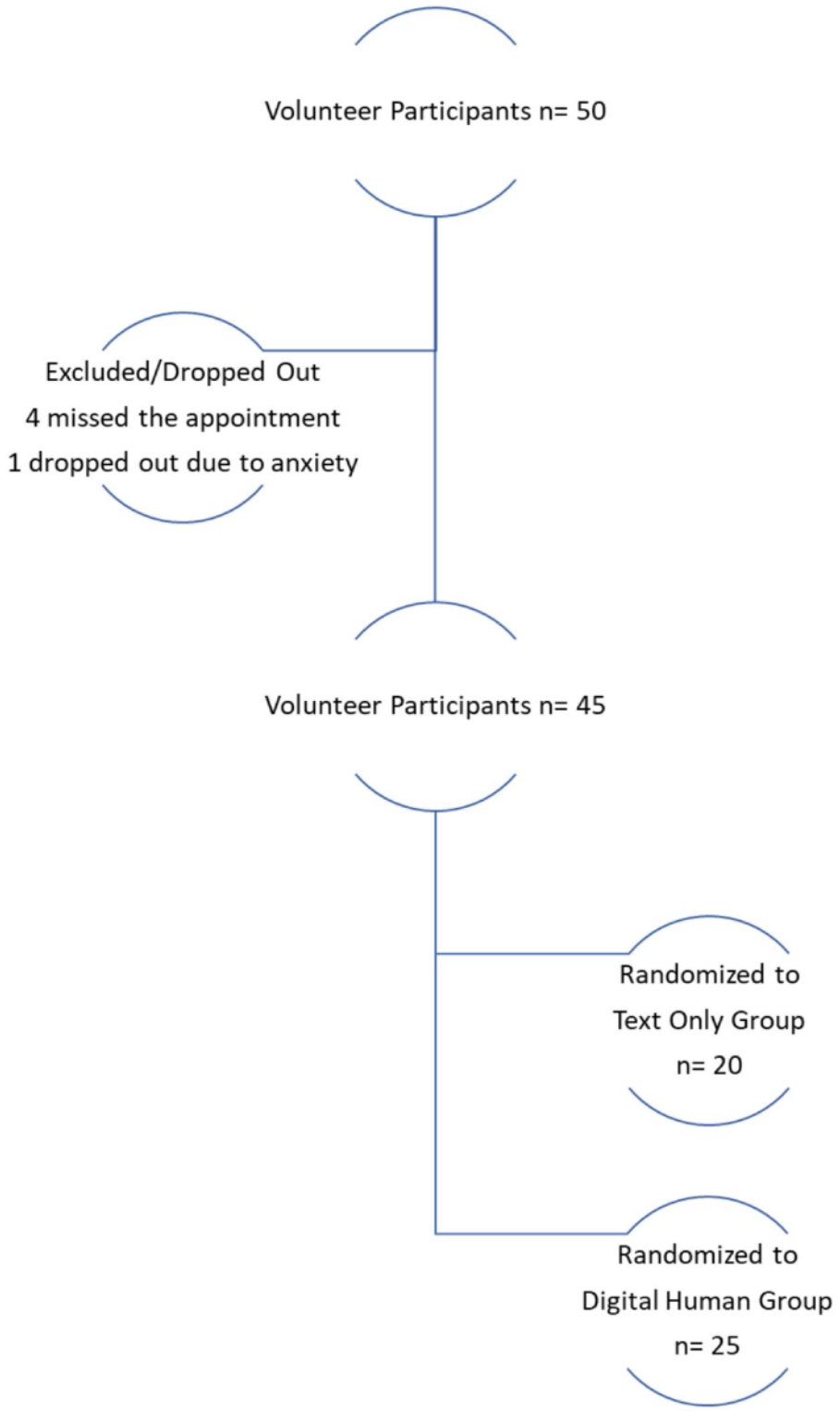
In this initial phase of system exploration, our focus was on evaluating the system's capabilities using volunteers who did not exhibit severe anxiety. As the system is still in its prototype stage, we exercised caution to avoid any potential exacerbation of symptoms in individuals with severe anxiety. Our recruitment announcement, disseminated through various social media channels associated with Sahlgrenska University's official account, specified that participants should be 18 years or older, free from any current mental health disorders, and willing to physically attend the testing facility in Gothenburg, Sweden.

Participants

Of the 50 individuals who initially volunteered, 5 participants (2 men and 3 women) opted out before providing their consent (Figure 2). Subsequently, 45 individuals attended the screening at the test facility. Each participant was required to provide informed consent before undergoing the Generalized Anxiety Disorder (GAD-7) scale assessment for anxiety symptoms. Those scoring 14 or higher on the GAD-7 were excluded from the study (Figure 2). Eligible participants were then randomly assigned to one of two groups: (1) engaging in text-based conversations with the text-only BETSY or (2) participating in voice-based interactions with the digital human (Figure 2).

XSL•FO
RenderX

Figure 2. Flow chart of participation and exclusion.



The randomization process was conducted with strict double-blind procedures overseen by an independent researcher not affiliated with this project and facilitated by an automated randomization system, ensuring the impartiality of the allocation.

Prechat Procedure

The experiments were performed during the COVID-19 pandemic (June 2021-November 2021). Due to safety precautions, the participants were greeted by a tester wearing protective gear, including a surgical R-II mask, gloves, a face visor, and hospital scrubs. Protective gear was also offered to participants upon their arrival. Participants were then placed in a sanitized room equipped with a screen, which underwent thorough sanitization with medical-grade disinfectants and a sterilizing UV lamp before and after each participant's session.

Before starting the chat with BETSY, participants were outfitted with a mobile dry-sensor EEG device to record their brain wave activity. Additionally, their blood pressure and pulse were

XSL•FO
RenderX

recorded on the left arm after a 5-minute seated rest. Systolic and diastolic blood pressure were measured using a digital sphygmomanometer, and the pulse was monitored with a pulse oximeter.

Despite relatively relaxed COVID-19 restrictions during the testing period, the tester opted not to be physically present in the room to minimize any potential risk of contagion. Each participant sat alone in the room, with the tester observing remotely via a nonrecordable streaming camera. This camera served to facilitate real-time communication and allowed the tester to monitor the participant's reactions and anticipate any need for assistance. The participants were made aware of this procedure.

Following the measurement of biometric data (blood pressure and pulse), participants were instructed to complete a questionnaire. This questionnaire covered their prior experiences with mental health chatbots as well as their demographic information, including sex, occupation, and marital status. Additionally, participants rated their overall well-being on a visual analog scale ranging from 1 (not good at all) to 10 (feeling excellent) before starting their session with BETSY. Participants were given instructions by the tester along with an accompanying sheet that provided potential chat scenarios and specified the topics within BETSY's scope. Each participant had a maximum of 30 minutes to engage with the chatbot version they were assigned to.

Chat Session Protocol and EEG Data Collection Process

EEG recording commenced simultaneously with the participant's initiation of their session with BETSY. We used a dry-sensor mobile EEG system, specifically the MUSE headband from Interaxon, which incorporates 7 sensors. These sensors include 3 frontal reference sensors and active sensors situated at Fp1, Fp2, Tp9, and Tp10.

The MUSE headband was seamlessly connected to a smartphone via Bluetooth and data collection was facilitated through the Mind Monitor app on an Android smartphone. It is noteworthy that this app neither necessitates user registration nor collects data that can identify or pinpoint individual users or their locations. Consequently, the data were recorded in an anonymous fashion and stored as a CSV file within the smartphone's document section.

Upon placing the EEG headband on the participant, they were requested to close their eyes to enable the Mind Monitor to perform calibration. After calibration was successfully established, the participant was left alone in the room to initiate a conversation with BETSY. Importantly, the EEG recording remained active throughout the entire conversation and was terminated when the participant indicated they had concluded their interaction with the chatbot.

Postchat Procedure

Upon reaching a point of satisfaction with the conversation or upon the completion of their allocated chat time, participants were directed to complete supplementary questionnaires and scales. Participants were administered the SUS-10, developed by Lewis and Sauro [26]. The SUS-10 calculates an average score derived from a 10-item questionnaire with response options ranging from 0 to 4, resulting in a total score between 0 and 100, as outlined by Bangor et al [31].

Participants were also presented with multiple-choice questions regarding their emotional state during the chat. Additionally, in the digital human group, participants were instructed to fill out the Standardized Questionnaires for Voice Interaction Design Short Version (SUISQ-MR). SUSIQ is a questionnaire tailored to assessing critical usability attributes of Interactive Voice Response, as outlined by Lewis and Sauro [26]. The original scale comprises 25 items categorized into 4 factors: user goal orientation, customer service behaviors, speech characteristics, and verbosity. SUISQ-MR, a shortened version, encompasses 9 items rated on a 7-point Likert scale, ranging from "strongly disagree" to "strongly agree." Higher scores on this scale indicate a more favorable assessment of the system's usability [24].

Furthermore, participants were provided with an open-ended questionnaire to gather their suggestions and insights regarding their session experience. It should be noted that qualitative data from this survey will be reported separately.

EEG Monitoring and Analysis

The monitoring of EEG activity entails the use of the Mind Monitor app, which captures and visually represents EEG brain wave data. The quantification of absolute brain wave values is predicated on the computation of absolute band powers. These powers are derived from the logarithm of the power spectral density calculated from the EEG data for each channel.

The frequency spectrum categories used for this analysis encompass the following bandwidths: δ (1-4 Hz), θ (4-8 Hz), α (7.5-13 Hz), β (13-30 Hz), and γ (30-44 Hz). Notably, the EEG power spectral density values acquired from the sensors typically fall within the {–1:+1} range, which is subsequently transformed into a more intelligible {0:100} range for text-based display purposes.

Subsequently, the collected EEG data underwent an analytical process facilitated by the Mind Monitor online graphing tool. Within this tool, the values are presented as average (dB) per session. It is imperative to note that, in the context of this study, we exclusively used absolute data for our analyses (information sourced from Mindmonitor.com and Choosemuse.com).

Statistical Analysis

All data were entered and processed in SPSS Statistics (version 28.0.1.1; IBM Corp). For group differences, means analysis was used using Pearson χ^2 asymptotic significance (2-sided) set at .05 as the significance level. For continuous outcome variables such as SUS-10, SUISQ-MR, brain wave activity, positivity, and GAD-7, linear regression analyses were used. The data were tested for kurtosis and skewness. Based on the results, *t* tests were performed. All results were analyzed according to group.

XSL•FO
RenderX

Ethical Considerations

All ethical decisions were guided by the Declaration of Helsinki and its subsequent amendments. The study protocol was reviewed and approved by the central ethical review board; Etikprövningsmyndigheten, Sweden (DRN 2021-02771). All precautions were taken in order to avoid any possible contagion from COVID-19. Due to the nature of the prototype, no patients were used in this initial examination of the chatbot in order to ensure that vulnerable participants would not be negatively affected by errors and flaws that might be present in a prototype-stage system.

Results

Characteristics of Participants

There were no statistically significant differences in the demographic variables between the digital human and text-only groups (Table 1). No participants were excluded due to a high GAD-7 score. The age of the participants ranged from 24 to 68 years, and as only 12 participants registered their age, this variable was consequently excluded from more advanced analyses (Table 1).

Table 1. Demographic characteristics of the study population.

Characteristic	Text-only, n (%)	Digital human, n (%)	*P* value[a]
Sex			.62
Male	8 (40)	10 (40)	
Female	12 (60)	15 (60)	
Marital status			.49
Married	11 (55)	11 (44)	
Single	5 (25)	3 (12)	
Divorced	1 (5)	2 (8)	
Domestic partnership	2 (10)	7 (28)	
Other	1 (5)	2 (8)	
Educational level			.74
High school/trade school	1 (5)	3 (12)	
Bachelor's	7 (35)	8 (32)	
Master's	7 (35)	11 (44)	
PhD	3 (15)	1 (4)	
Other/higher than master's or PhD	2 (10)	3 (8)	
Occupation			.30
Sick leave/sick leave part-time	0 (0)	1 (4)	
Working part-time	0 (0)	2 (8)	
Working full time	18 (90)	19 (76)	
Student	2 (10)	1 (4)	
Retired	0 (0)	2 (8)	
Housing			.35
Living alone	4 (20)	2 (8)	
Cohabitation	16 (80)	23 (92)	

[a]Pearson χ^2 test.

Comparison Between Digital Human and Text-Only Chatbots

When comparing self-reported emotional states between the digital human and the text-only chatbot groups, it was observed that participants using the digital human exhibited a notably higher propensity to report feelings of nervousness versus the text-only chatbot group (Table 2). The mean GAD-7 score for the text-only chatbot group was 2.32 (SD 2.52) compared with 2.80 (SD 2.60) for the digital human chatbot group, with no statistically significant difference between the groups.

XSL•FO
RenderX

Table 2. Self-reported prior therapy experience, emotions, biometrics, and electroencephalography.

	Text-only	Digital human	*P* value[a]
Therapy, n (%)			.47
Yes	2 (10)	4 (16)	
No	18 (90)	20 (80)	
Do not remember	0	1 (4)	
"Have you talked to a chatbot about mental health before?", n (%)			.84
Yes	1 (5)	1 (4)	
No	19 (95)	24 (96)	
Do not remember	0	0	
GAD-7[b] score, mean (SD)[a]	2.3 (2.5)	2.8 (2.6)	.56
Positivity toward chatbot, mean (SD)[c]	7.1 (2.1)	7.5 (2.1)	.69
"Do you feel closeness to BETSY[d]?", n (%)			.46
Yes	7 (35)	11 (45.8)	
No	13 (65)	13 (54.2)	
"Did you feel relaxed?", n (%)			.23
Yes or sometimes	17 (89.5)	17 (73.9)	
No	2 (10.5)	6 (26.1)	
"Did you feel nervous?", n (%)			.02
Yes or sometimes	0	6 (26.1)	
No	19 (100)	17 (73.9)	
"Did you feel sad?", n (%)			.1
Yes or sometimes	0	3 (13.0)	
No	19 (100)	20 (87.0)	
"Did you feel annoyance?", n (%)			.8
Yes or sometimes	9 (47.4)	10 (43.5)	
No	10 (52.6)	13 (56.5)	
VAS-W[e] presession, mean (SD)[f]	8.8 (1.32)	8.4 (1.41)	.33
VAS-W postsession, mean (SD)[f]	8.8 (1.23)	8.3 (1.27)	.14
Pulse presession, mean (SD)	72.2 (10.7)	71.6 (10.9)	.77
Pulse postsession, mean (SD)	68.5 (8.8)	70.2 (11.1)	.58
Average δ wave activity, mean (SD)	114 (30)	97 (25)	.06
Average θ wave activity, mean (SD)	86 (23)	74 (21)	.08
Average α wave activity, mean (SD)	97 (27)	82 (24)	.03
Average β wave activity, mean (SD)	81 (17)	76 (20)	.34
Average γ wave activity, mean (SD)	65 (15)	66 (21)	.98
SUS-10[g], mean (SD)	74.82 (10)	64.80 (14)	.01
SUISQ-MR[h], mean (range)	N/A[i]	4.92 (2.83-6.75)	N/A

[a]Pearson χ^2 test for categorical variables and ANOVA for continuous variables.

[b]GAD-7: Generalized Anxiety Disorder Scale.

[c]Participants were asked to what extent they felt positive about talking to BETSY about mental health with scores ranging from 1 (not positive at all) to 10 (very positive).

[d]BETSY: Behavior, Emotion, Therapy System, and You.

XSL•FO
RenderX

[e]VAS-W: Visual Analogue Scale for Well-Being.

[f]Range from 1 (not feeling well at all) to 10 (feeling very good).

[g]SUS-10: System Usability Scale.

[h]SUISQ-MR: Standardized Questionnaires for Voice Interaction Design Short Version.

[i]N/A: not applicable.

Conversely, the evaluation of system usability as gauged by SUS-10 showed a significant (P=.01) difference between the groups. Notably, the mean SUS-10 score was higher in the text-only chatbot group at 75.34 (SD 10.01; range 57-90) compared with the digital human group at 64.80 (SD 14.14; range 40-90). In addition, the digital human group underwent assessment using SUISQ-MR: BETSY had a mean score of 4.92 (SD 0.83; range 2.83-6.75), as depicted in Table 2, which is indicative of a commendable level of usability for BETSY's voice interface in accordance with the framework presented by Lewis [24].

Biometric Measures

There were no statistically significant distinctions for mean values of blood pressure or pulse between the groups either at baseline or following exposure to the interventions. Specifically, the mean pulse rate showed no discernible variations between the groups both before and after exposure, reflecting consistent values across the groups on average (data not shown).

The EEG signals collected during the study exhibited suboptimal quality, which was primarily attributed to participant movement and signal acquisition sensitivity. These challenges occasionally disrupted signal continuity during the sessions. Nonetheless, the data yielded adequate information to calculate mean values pertaining to δ, θ, α, β, and γ frequency bands, as facilitated by the web-based graphing module within the MindMonitor's platform. Only 1 significant difference was found in terms of means: the average α was significantly higher in the text-only group (Table 2).

System Usability Scale and Outcomes

Predictors of SUS-10 usability were used as a dependent variable in linear regression analysis and matched against biometric and subjective variables. Each variable was independently analyzed in a model together with SUS-10 as the dependent variable. Analysis showed that there was a significant positive relationship between average α and θ wave activity and SUS-10 in the chat-only group. A significant positive relationship was seen between SUISQ-MR scores and SUS-10 (Table 3).

XSL•FO
RenderX

Table 3. Linear regression analysis between usability and biometric variables.

	Unstandardized coefficients		*P* value[a]
	β	SE	
SUS-10[b] × positivity			
Text (n=19)	1.82	1.02	.09
Voice (n=24)	1.313	1.361	.35
SUS-10 × average δ wave activity			
Text (n=17)	0.153	0.08	.07
Voice (n=23)	0.062	0.12	.61
SUS-10 × average θ wave activity			
Text (n=17)	0.212	0.1	.05
Voice (n=23)	0.083	0.146	.57
SUS-10 × average α wave activity			
Text (n=17)	0.196	0.083	.03
Voice (n=23)	0.054	0.124	.67
SUS-10 × average β wave activity			
Text (n=17)	0.251	0.143	.10
Voice (n=23)	0.03	0.152	.85
SUS-10 × average γ wave activity			
Text (n=17)	0.148	0.177	.42
Voice (n=23)	–0.017	0.146	.91
SUS-10 × SUISQ-MR[c] (n=24)	8.100	2.976	.01

[a]Pearson χ^2 test.

[b]SUS-10: System Usability Scale.

[c]SUISQ-MR: Standardized Questionnaires for Voice Interaction Design Short Version.

Self-Reported Feelings and Gender

Furthermore, our investigation sought to discern whether significant gender disparities existed in terms of self-reported emotions. Notably, we observed a significant difference between men and women, with men exhibiting a notably lower tendency to report feeling annoyed by BETSY in contrast to women. No other statistically significant distinctions were identified (Table 4).

XSL•FO
RenderX

Table 4. Sex difference in emotional expression toward BETSY (Behavior, Emotion, Therapy System, and You).

Self-reported feeling	Men, n (%)	Women, n (%)	*P* value[a]
Did you feel annoyed?			
Chat			.03
Yes	1 (11)	8 (67)	
No	6 (88)	4 (33)	
Voice			.03
Yes	1 (12.5)	9 (60)	
No	7 (84.5)	6 (40)	
Did you feel relaxed?			
Chat			.68
Yes	6 (86)	11 (92)	
No	1 (14)	1 (8)	
Voice			.29
Yes	7 (87.5)	10 (67)	
No	1 (12.5)	5 (33)	
Did you feel closeness or connection to BETSY[b]?			
Chat			.25
Yes	4 (50)	3 (25)	
No	4 (50)	9 (75)	
Voice			.63
Yes	4 (40)	7 (50)	
No	6 (60)	7 (50)	
Did you feel nervous?			
Chat			
Yes	N/A[c]	N/A	
No	7 (100)	12 (100)	
Voice			.93
Yes	2 (25)	4 (27)	
No	6 (75)	11 (73)	
Did you feel sadness?			
Chat			
Yes	N/A	N/A	
No	7 (100)	12 (100)	
Voice			.17
Yes	0	3 (20)	
No	8 (100)	12 (80)	

[a]Pearson χ^2 test.

[b]BETSY: Behavior, Emotion, Therapy System, and You.

[c]N/A: not applicable.

An analysis of feelings of closeness and positivity toward chatbot conversations was undertaken to explore differences between men and women. In mean score analyses, the results showed that men were significantly more positive toward talking to BETSY prior to the session: 8.16 (SD 1.50) for men and 6.81 (SD 2.30) for women (*P*=.34). Conversely, there were no discernible gender-based differences concerning feelings of closeness during chatbot interactions.

XSL•FO
RenderX

Discussion

Principal Findings

This study explored how a digital human versus text-only chatbot interface affected usability and user experience in healthy participants. We also examined how chatbot-generated conversations on mental health affected self-reported feelings and biometrics. The overall sample was small and, thus, should not act as a point of reference for generalization. This study was, however, not smaller than the average study in the investigative field of mental health chatbots [9,10,12,37-41].

While the text-only system scored higher on usability, both versions of the chatbot scored average or above average with respect to overall usability [31]. The mean text-only chatbot SUS-10 score of 75.34 falls between the threshold of good (a score of 70) and excellent (a score of 80 and above) [29-31]. However, the score for the digital human (64.8) indicates that the system is perceived to be usable, but has room for improvement. Usability can be affected by many factors such as user interface design, content layout, and overall user experience [42,43].

The digital human score indicates that there may be areas for improvement in terms of all of the aforementioned aspects. It should also be noted that the SUS-10 scale does not measure a specific feature or aspect of system design, but instead provides an overall assessment of user experience [31]. Using more elaborate scales that cover more dimensions across the system is more suitable for a more in-depth analysis of the usability of chatbots. It can also be noted that the range of scores was much higher for the text-only interface (lowest score for the text-only group was 57 and the equivalent for the voice-only chatbot was 40), which indicates much poorer usability.

Taking into account the specific usability of the digital human interface, usability was considered high with an average SUISQ-MR score of 4.92. This score indicates that the voice interaction design is likely to be perceived as intuitive and useful by users. A score of 4.92 falls within the range of 4.5-5.5, which has been classified as “very good” in previous studies [24]. In addition, higher scores on the SUISQ-MR have been associated with increased user engagement and task completion rates. Therefore, a score of 4.92 can be interpreted as an indication that the voice interaction design will likely provide positive user experiences [24].

Men were more likely to score higher on positivity and less likely to report feeling annoyed by BETSY independently of the interface, which is the opposite of other studies that indicate men are more likely to be annoyed or aggressive toward female avatars [44]. In a study by Luger and Sellen [45], the authors found that higher expectations of the system lead to a higher risk of disappointment and lower scores: this could possibly explain why female participants were more agitated as their expectations might have been higher [45], however, we have no data to explore this empirically in the frame of this study. Unlike other studies with similar designs and populations [40], we did not analyze the content of the conversations. The conversations between BETSY and the participants were deleted immediately after the session as the research question was geared toward usability and not the effect on the user’s own mental health status. Much like the results from Hearst and Tory [46], the interface was the focus of this investigation. Hearst and Tory showed that a well-designed conversation tamped the choice of interface. The interface played into the perception of usability only when the system failed to respond or create barriers to conversation. In our study, we used biometric data to explore feelings of relaxation or excitement/agitation while using BETSY. Despite EEG data collection not being optimal, we were able to collect and compare some brain wave activity data in the study groups during the sessions. Even though the amplitude of brain wave activity can result in large intraindividual variation, the data were evenly distributed and there were no mean differences between the groups in our study. While we observed no significant association between scores of usability and β wave activity (more likely to be associated with frustration, agitation, or perhaps excitement), we did observe brain wave activity that is typically associated with relaxed states of mind [18-21] and this had a positive linear relationship with SUS-10 score. This indicated that the chat-only group was either more relaxed or less aroused (or both). The explanation could lie in the combination of the small sample and the fact that more individuals in the voice-only chat group reported feeling nervousness, a feeling that generally elicits higher brain wave activity and less relaxation [20,47]. With the low quality of data, combined with a limited sample, it is hard to draw any generalizable conclusions from the biometric data.

Feelings of closeness did not differ between the 2 interfaces and seem not to have been affected by the presence of anthropomorphic features. When gender was explored as a factor, there was no significant difference to what extent men and women reported feeling close to BETSY in the respective assigned interfaces. Due to the small sample size in our study, it was not possible to perform further and more elaborate designs looking at mediation of other demographic or biometric factors in a reliable way.

When devising chatbots for mental health, this study indicates that a mixed approach might be the best course of action, allowing the user to choose a preferred way of interacting with the chatbot.

Limitations

This study consisted of healthy volunteers. It is good to keep in mind that mental health issues can affect some parts of cognitive performance [48] and, thus, usability may not be equally perceived by a person in a state of emotional distress and a healthy volunteer. Further investigation and collaboration are needed in future studies to capture the usability aspects of individuals who are in an active state of distress.

The results of this study suggest that overall usability seems to be perceived as higher for the text-only chatbot interface and no significant emotional boost was present with the addition of anthropomorphic features to a digital human chatbot.

Further studies which include a larger sample of participants as well as participants who experience mild to moderate anxiety are needed to explore and further evaluate the research question

XSL•FO
RenderX

posed in this paper. In this study, the age range was limited, and the variable was incomplete. In future studies, we will strive to include more young adults and adults older than 60 years.

Large language models and application programming interface models were not available at the time this chatbot was constructed and neither were Metahuman creator or more advanced voice-cloning or voice-generating options, which would have significantly improved the anthropomorphic features of the digital human. The first iteration of the generative pretrained transformer was not available to the public and the generative pretrained transformer-3 application programming interface had a limited release during the development of this project: it was not available to our team until a year after the project was completed. With large language models, repetitiveness and limitations in terms of variability of answers would have most likely been avoided; however, the aim of our investigation was not the general effect of the content but rather the perception of text- versus voice-driven interfaces.

Conclusions

In conclusion, the text-only chatbot was perceived as more user-friendly in terms of usability indicators for SUS-10. However, both the digital human and text-only interfaces scored average or above average in comparison to other studies performed on mental health chatbots. Although biometric data did not differ significantly, we saw significant gender differences in terms of prechat positivity and postchat annoyance, which is contrary to other studies. Male participants in our study were more likely to report higher prechat positivity toward BETSY and report less irritation postchat. SUISQ-MR also indicated that BETSY's overall usability and voice were highly ranked compared with other studies, indicating that there is great promise for mental health chatbots independently of the chosen user interface.

Acknowledgments

The authors would like to thank the participants of this study for their time and contribution to valuable knowledge. The authors also especially thank Mathias Alvidius, head of department, for contributing significantly to the possibilities of the creation of this research project. The authors would also like to express sincere gratitude to Linn Dahlén Ölander for her help in recruiting participants. SS received support from the Söderström König Foundation (SLS-974247), Stiftelsen Psykiatriska Forskningsfonden for this project. Parts of this project were funded by Innovationsfonden VGR (project VGRINN-936800) and Västra Götalandsregionen. The authors would also like to thank Peter Todd (Tajut Ltd) for third-party editorial assistance in the drafting of this manuscript, for which he received financial compensation.

Conflicts of Interest

SS receives fees for scientific consultation from Mindforce.

Multimedia Appendix 1
CONSORT-EHEALTH checklist (V 1.6.1).
[PDF File (Adobe PDF File), 107 KB - humanfactors_v11i1e54581_app1.pdf]

References

1. Nadkarni PM, Ohno-Machado L, Chapman WW. Natural language processing: an introduction. J Am Med Inform Assoc 2011;18(5):544-551 [FREE Full text] [doi: 10.1136/amiajnl-2011-000464] [Medline: 21846786]
2. Weizenbaum J. ELIZA—a computer program for the study of natural language communication between man and machine. Commun ACM 1966;9(1):36-45 [FREE Full text] [doi: 10.1145/365153.365168]
3. Shum H, He X, Li D. From Eliza to XiaoIce: challenges and opportunities with social chatbots. Front Inf Technol Electronic Eng 2018;19(1):10-26 [FREE Full text] [doi: 10.1631/fitee.1700826]
4. Lim SM, Shiau CWC, Cheng LJ, Lau Y. Chatbot-delivered psychotherapy for adults with depressive and anxiety symptoms: a systematic review and meta-regression. Behav Ther 2022;53(2):334-347 [FREE Full text] [doi: 10.1016/j.beth.2021.09.007] [Medline: 35227408]
5. Radziwill NM, Benton MC. Evaluating quality of chatbots and intelligent conversational agents. ArXiv. Preprint posted online on April 15 2017 [FREE Full text]
6. Dale R. GPT-3: what's it good for? Nat Lang Eng 2020;27(1):113-118 [FREE Full text] [doi: 10.1017/S1351324920000601]
7. Homolak J. Opportunities and risks of ChatGPT in medicine, science, and academic publishing: a modern Promethean dilemma. Croat Med J 2023;64(1):1-3 [FREE Full text] [doi: 10.3325/cmj.2023.64.1] [Medline: 36864812]
8. Caldarini G, Jaf S, McGarry K. A literature survey of recent advances in chatbots. Information 2022;13(1):41. [doi: 10.3390/info13010041]
9. Cameron G, Cameron D, Megaw G, Bond R, Mulvenna M, O'Neill S, et al. Towards a chatbot for digital counselling. 2017 Presented at: Proceedings of the 31st International BCS Human Computer Interaction Conference (HCI 2017) (HCI); July 3-6, 2017; Sunderland, UK URL: https://www.scienceopen.com/hosted-document?doi=10.14236/ewic/HCI2017.24 [doi: 10.14236/ewic/hci2017.24]

XSL•FO
RenderX

10. Denecke K, Vaaheesan S, Arulnathan A. A mental health chatbot for regulating emotions (SERMO) - concept and usability test. IEEE Trans Emerg Topics Comput 2020;9(3):1170-1182 [FREE Full text] [doi: 10.1109/tetc.2020.2974478]
11. Torous J, Bucci S, Bell IH, Kessing LV, Faurholt-Jepsen M, Whelan P, et al. The growing field of digital psychiatry: current evidence and the future of apps, social media, chatbots, and virtual reality. World Psychiatry 2021;20(3):318-335 [FREE Full text] [doi: 10.1002/wps.20883] [Medline: 34505369]
12. Oh J, Jang S, Kim H, Kim JJ. Efficacy of mobile app-based interactive cognitive behavioral therapy using a chatbot for panic disorder. Int J Med Inform 2020;140:104171 [FREE Full text] [doi: 10.1016/j.ijmedinf.2020.104171] [Medline: 32446158]
13. Kaywan P, Ahmed K, Ibaida A, Miao Y, Gu B. Early detection of depression using a conversational AI bot: a non-clinical trial. PLoS One 2023;18(2):e0279743 [FREE Full text] [doi: 10.1371/journal.pone.0279743] [Medline: 36735701]
14. Islam MN, Khan SR, Islam NN, Rownok RA, Zaman SR, Zaman SR. A mobile application for mental health care during COVID-19 pandemic: development and usability evaluation with System Usability Scale. : Springer; 2021 Presented at: Computational Intelligence in Information Systems: Proceedings of the Computational Intelligence in Information Systems Conference (CIIS 2020); January 25–27, 2021; Brunei Darussalam p. 33-42. [doi: 10.1007/978-3-030-68133-3_4]
15. Heller B, Proctor M, Mah D, Jewell L, Cheung B. Freudbot: an investigation of chatbot technology in distance education. In: EdMedia+ Innovate Learning. Paris, France: Association for the Advancement of Computing in Education (AACE); 2005:3913-3918.
16. Kadariya D, Venkataramanan R, Yip HY, Kalra M, Thirunarayanan K, Sheth A. kBot: knowledge-enabled personalized chatbot for asthma self-management. Proc Int Conf Smart Comput SMARTCOMP 2019;2019:138-143 [FREE Full text] [doi: 10.1109/smartcomp.2019.00043] [Medline: 32832938]
17. Sagar M, Seymour M, Henderson A. Creating connection with autonomous facial animation. Commun ACM 2016;59(12):82-91 [FREE Full text] [doi: 10.1145/2950041]
18. Allen JJB, Reznik SJ. Frontal EEG asymmetry as a promising marker of depression vulnerability: summary and methodological considerations. Curr Opin Psychol 2015;4:93-97 [FREE Full text] [doi: 10.1016/j.copsyc.2014.12.017] [Medline: 26462291]
19. Hagemann D, Naumann E, Lürken A, Becker G, Maier S, Bartussek D. EEG asymmetry, dispositional mood and personality. Pers Individ Dif 1999;27(3):541-568 [FREE Full text] [doi: 10.1016/s0191-8869(98)00263-3]
20. Lomas T, Ivtzan I, Fu CHY. A systematic review of the neurophysiology of mindfulness on EEG oscillations. Neurosci Biobehav Rev 2015;57:401-410 [FREE Full text] [doi: 10.1016/j.neubiorev.2015.09.018] [Medline: 26441373]
21. Al-Nafjan A, Hosny M, Al-Ohali Y, Al-Wabil A. Review and classification of emotion recognition based on EEG brain-computer interface system research: a systematic review. Appl Sci 2017;7(12):1239. [doi: 10.3390/app7121239]
22. Bos AS, Pizzato MC, Vettori M, Donato LG, Soares PP, Fagundes JG, et al. Empirical evidence during the implementation of an educational chatbot with the electroencephalogram metric. Creat Educ 2020;11(11):2337-2345 [FREE Full text] [doi: 10.4236/ce.2020.1111171]
23. Ciechanowski L, Przegalinska A, Magnuski M, Gloor P. In the shades of the uncanny valley: an experimental study of human–chatbot interaction. Future Gener Comput Syst 2019;92:539-548 [FREE Full text] [doi: 10.1016/j.future.2018.01.055]
24. Lewis JR. Standardized questionnaires for voice interaction design. Voice Interact Des 2016;1(1):1-16 [FREE Full text] [doi: 10.1016/b978-1-55860-768-2.x5000-2]
25. Lewis JR. The System Usability Scale: past, present, and future. Int J Hum Comput Interact 2018;34(7):577-590 [FREE Full text] [doi: 10.1080/10447318.2018.1455307]
26. Lewis JR, Sauro J. The factor structure of the System Usability Scale. 2009 Presented at: Human Centered Design: First International Conference, HCD 2009, held as part of HCI International; July 19-24, 2009; San Diego, CA. [doi: 10.1007/978-3-642-02806-9_12]
27. Lewis JR, Hardzinski ML. Investigating the psychometric properties of the speech user interface service quality questionnaire. Int J Speech Technol 2015;18(3):479-487 [FREE Full text] [doi: 10.1007/s10772-015-9289-1]
28. Peres SC, Pham T, Phillips R. Validation of the System Usability Scale (SUS): SUS in the wild. Proc Hum Factors Ergon Soc Annu Meet 2013;57(1):192-196 [FREE Full text] [doi: 10.1177/1541931213571043]
29. Grier RA, Bangor A, Kortum P, Peres SC. The System Usability Scale: beyond standard usability testing. Proc Hum Factors Ergon Soc Annu Meet 2013;57(1):187-191 [FREE Full text] [doi: 10.1177/1541931213571042]
30. Kortum PT, Bangor A. Usability ratings for everyday products measured with the System Usability Scale. Int J Hum Comput Interact 2013;29(2):67-76 [FREE Full text] [doi: 10.1080/10447318.2012.681221]
31. Bangor A, Kortum PT, Miller JT. An empirical evaluation of the System Usability Scale. Int J Hum Comput Interact 2008;24(6):574-594 [FREE Full text] [doi: 10.1080/10447310802205776]
32. Bennion MR, Hardy GE, Moore RK, Kellett S, Millings A. Usability, acceptability, and effectiveness of web-based conversational agents to facilitate problem solving in older adults: controlled study. J Med Internet Res 2020;22(5):e16794 [FREE Full text] [doi: 10.2196/16794] [Medline: 32384055]
33. Ly KH, Ly AM, Andersson G. A fully automated conversational agent for promoting mental well-being: a pilot RCT using mixed methods. Internet Interv 2017;10:39-46 [FREE Full text] [doi: 10.1016/j.invent.2017.10.002] [Medline: 30135751]

XSL•FO
RenderX

34. Ta V, Griffith C, Boatfield C, Wang X, Civitello M, Bader H, et al. User experiences of social support from companion chatbots in everyday contexts: thematic analysis. J Med Internet Res 2020;22(3):e16235 [FREE Full text] [doi: 10.2196/16235] [Medline: 32141837]
35. van Wezel MMC, Croes EAJ, Antheunis ML. “I’m here for you”: can social chatbots truly support their users? a literature review. 2020 Presented at: Chatbot Research and Design: Proceedings of the 4th International Workshop, CONVERSATIONS 2020; November 23-24, 2020; Virtual Event p. 96-113. [doi: 10.1007/978-3-030-68288-0_7]
36. Crolic C, Thomaz F, Hadi R, Stephen AT. Blame the bot: anthropomorphism and anger in customer–chatbot interactions. J Mark 2021;86(1):132-148 [FREE Full text] [doi: 10.1177/00222429211045687]
37. Greer S, Ramo D, Chang YJ, Fu M, Moskowitz J, Haritatos J. Use of the chatbot "Vivibot" to deliver positive psychology skills and promote well-being among young people after cancer treatment: randomized controlled feasibility trial. JMIR Mhealth Uhealth 2019;7(10):e15018 [FREE Full text] [doi: 10.2196/15018] [Medline: 31674920]
38. Jang S, Kim JJ, Kim SJ, Hong J, Kim S, Kim E. Mobile app-based chatbot to deliver cognitive behavioral therapy and psychoeducation for adults with attention deficit: a development and feasibility/usability study. Int J Med Inform 2021;150:104440 [FREE Full text] [doi: 10.1016/j.ijmedinf.2021.104440] [Medline: 33799055]
39. Loveys K, Sagar M, Pickering I, Broadbent E. A digital human for delivering a remote loneliness and stress intervention to at-risk younger and older adults during the COVID-19 pandemic: randomized pilot trial. JMIR Ment Health 2021;8(11):e31586 [FREE Full text] [doi: 10.2196/31586] [Medline: 34596572]
40. Ruane E, Farrell S, Ventresque A. User perception of text-based chatbot personality. 2020 Presented at: Chatbot Research and Design: Proceedings of the 4th International Workshop, CONVERSATIONS 2020; November 23-24, 2020; Virtual event p. 32-47. [doi: 10.1007/978-3-030-68288-0_3]
41. Yen C, Chiang MC. Trust me, if you can: a study on the factors that influence consumers’ purchase intention triggered by chatbots based on brain image evidence and self-reported assessments. Behav Inf Technol 2020;40(11):1177-1194 [FREE Full text] [doi: 10.1080/0144929x.2020.1743362]
42. Cheng Y, Jiang H. How do AI-driven chatbots impact user experience? examining gratifications, perceived privacy risk, satisfaction, loyalty, and continued use. J Broadcast Electron Media 2020;64(4):592-614 [FREE Full text] [doi: 10.1080/08838151.2020.1834296]
43. Smestad TL, Volden F. Chatbot personalities matters: improving the user experience of chatbot interfaces. 2018 Presented at: Internet Science: 5th International Conference, INSCI 2018; October 24–26, 2018; St. Petersburg, Russia p. 170-181. [doi: 10.1007/978-3-030-17705-8_15]
44. Silvervarg A, Raukola K, Haake M, Gulz A. The effect of visual gender on abuse in conversation with ECAs. 2012 Presented at: Intelligent Virtual Agents: 12th International Conference, IVA 2012; September 12-14, 2012; Santa Cruz, CA p. 153-160. [doi: 10.1007/978-3-642-33197-8_16]
45. Luger E, Sellen A. "Like having a really bad PA": the gulf between user expectation and experience of conversational agents. 2016 Presented at: CHI'16: CHI Conference on Human Factors in Computing Systems; May 7-12, 2016; San Jose, CA p. 5286-5297 URL: https://dl.acm.org/doi/10.1145/2858036.2858288 [doi: 10.1145/2858036.2858288]
46. Hearst M, Tory M. Would you like a chart with that? incorporating visualizations into conversational interfaces. 2019 Presented at: 2019 IEEE Visualization Conference (VIS); October 20-25, 2019; Vancouver, BC p. 1-5. [doi: 10.1109/visual.2019.8933766]
47. Díaz H, Cid FM, Otárola J, Rojas R, Alarcón O, Cañete L. EEG Beta band frequency domain evaluation for assessing stress and anxiety in resting, eyes closed, basal conditions. Procedia Comput Sci 2019;162:974-981 [FREE Full text] [doi: 10.1016/j.procs.2019.12.075]
48. Eysenck MW, Derakshan N, Santos R, Calvo MG. Anxiety and cognitive performance: attentional control theory. Emotion 2007;7(2):336-353. [doi: 10.1037/1528-3542.7.2.336] [Medline: 17516812]

Abbreviations

BETSY: Behavior, Emotion, Therapy System, and You
EEG: electroencephalography
GAD-7: Generalized Anxiety Disorder-7
SUS-10: System Usability Scale-10
SUISQ-MR: Standardized Questionnaires for Voice Interaction Design Short Version

XSL•FO
RenderX

Edited by A Kushniruk; submitted 15.11.23; peer-reviewed by R Weeks, R Pryss; comments to author 22.12.23; revised version received 27.01.24; accepted 18.02.24; published 29.04.24.

Please cite as:
Thunström AO, Carlsen HK, Ali L, Larson T, Hellström A, Steingrimsson S
Usability Comparison Among Healthy Participants of an Anthropomorphic Digital Human and a Text-Based Chatbot as a Responder to Questions on Mental Health: Randomized Controlled Trial
JMIR Hum Factors 2024;11:e54581
URL: https://humanfactors.jmir.org/2024/1/e54581
doi:10.2196/54581
PMID:38683664

©Almira Osmanovic Thunström, Hanne Krage Carlsen, Lilas Ali, Tomas Larson, Andreas Hellström, Steinn Steingrimsson. Originally published in JMIR Human Factors (https://humanfactors.jmir.org), 29.04.2024. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on https://humanfactors.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Identifying Factors of User Acceptance of a Drone-Based Medication Delivery: User-Centered Design Approach

Franziska Fink[1,2*], Dr rer nat; Ivonne Kalter[1,2*], MA; Jenny-Victoria Steindorff[1,2], MSc; Hans Konrad Helmbold[3], MSc; Denny Paulicke[1,2,4], Prof Dr; Patrick Jahn[1,2], Prof Dr

[1]Translation Region for Digitalised Healthcare, Department of Internal Medicine, Faculty of Medicine, University Medicine, Martin-Luther-University Halle-Wittenberg, Halle (Saale), Germany

[2]Health Service Research Working Group | Acute Care, Department of Internal Medicine, Faculty of Medicine, University Medicine, Martin-Luther-University Halle-Wittenberg, Halle (Saale), Germany

[3]Department of Economics, Anhalt University of Applied Sciences, Bernburg, Germany

[4]Department of Medical Pedagogy, Akkon University of Human Sciences, Berlin, Germany

[*]these authors contributed equally

Corresponding Author:
Patrick Jahn, Prof Dr
Translation Region for Digitalised Healthcare
Department of Internal Medicine, Faculty of Medicine, University Medicine
Martin-Luther-University Halle-Wittenberg
Magdeburger Str. 12
Halle (Saale), 06112
Germany
Phone: 49 3455575493
Email: patrick.jahn@uk-halle.de

Abstract

Background: The use of drones in the health care sector is increasingly being discussed against the background of the aging population and the growing shortage of skilled workers. In particular, the use of drones to provide medication in rural areas could bring advantages for the care of people with and without a need for care. However, there are hardly any data available that focus on the interaction between humans and drones.

Objective: This study aims to disclose and analyze factors associated with user acceptance of drone-based medication delivery to derive practice-relevant guidance points for participatory technology development (for apps and drones).

Methods: A controlled mixed methods study was conducted that supports the technical development process of an app design for drone-assisted drug delivery based on a participatory research design. For the quantitative analysis, established and standardized survey instruments to capture technology acceptance, such as the System Usability Scale; Technology Usage Inventory (TUI); and the Motivation, Engagement, and Thriving in User Experience model, were used. To avoid possible biasing effects from a continuous user development (eg, response shifts and learning effects), an ad hoc group was formed at each of the 3 iterative development steps and was subsequently compared with the consisting core group, which went through all 3 iterations.

Results: The study found a positive correlation between the usability of a pharmacy drone app and participants' willingness to use it (r=0.833). Participants' perception of usefulness positively influenced their willingness to use the app (r=0.487; TUI). Skepticism had a negative impact on perceived usability and willingness to use it (r=−0.542; System Usability Scale and r=−0.446; TUI). The study found that usefulness, skepticism, and curiosity explained most of the intention to use the app ($F_{3,17}$=21.12; P<.001; R^2=0.788; adjusted R^2=0.751). The core group showed higher ratings on the intention to use the pharmacy drone app than the ad hoc groups. Results of the 2-tailed t tests showed a higher rating on usability for the third iteration of the core group compared with the first iteration.

Conclusions: With the help of the participatory design, important aspects of acceptance could be revealed by the people involved in relation to drone-assisted drug delivery. For example, the length of time spent using the technology is an important factor for the intention to use the app. Technology-specific factors such as user-friendliness or curiosity are directly related to the use acceptance of the drone app. Results of this study showed that the more participants perceived their own competence in handling the app, the more they were willing to use the technology and the more they rated the app as usable.

XSL•FO
RenderX

(JMIR Hum Factors 2024;11:e51587) doi:10.2196/51587

KEYWORDS

human-drone interaction; medical supplies; participative research; user-centered design; technology acceptance

Introduction

Background

The health care system faces challenges such as a rural exodus, aging populations, and increasing shortages in the health care workforce; drones have the potential to increase the efficiency and capacity of the health care system [1]. The COVID-19 pandemic intensified the demand for new logistic solutions, such as fast and contactless delivery strategies [2]. It is also important to understand the attitude of the civilian population and public opinion on the use of drones [3]. In this vein, delivering medications with drones is the most identified application in health care [1,4]. There are a few studies showing that usability, lack of user skills and expertise, and negative perceptions affect user acceptance and hinder drone use [1,5,6]. Therefore, it is particularly important that all user groups are involved at an early stage. Furthermore, after a recent scoping review showed that there were no empirical studies on user acceptance of drone-based medication delivery [7], we could only find 1 study in Asia that investigated user acceptance among health care professionals in the delivery of drone-based medication [8]. The successful application of technology is predominantly determined by the type and extent of acceptance [9,10]. Acceptance in this context refers to the positive acceptance decision of an innovation by its users, which is described in the technology acceptance model (TAM; perceived usability, usefulness, immersion, and accessibility: TAM 1) [11,12]. It proposes that users tend to use a technology when they believe it will help them to perform a better job (perceived usefulness) and when they believe that the system can be handled without effort (perceived ease of use). These variables were found to correlate with the intention to use, wherein usefulness was substantially more strongly related to frequency of use than ease of use. Nevertheless, both are strong correlates of user acceptance and should not be ignored by designing and implementing successful technologies [12]. In other words, the greater the benefit of a technology and the simpler its usability, the more the users are willing to use the new system. However, some more variables that affect user acceptance such as social influences (subjective norm, image, and voluntariness), cognitive instrumental processes (job relevance, output quality, and result demonstrability; TAM 2), and psychological foundations (self-efficacy, external control, playfulness, anxiety, enjoyment, and usability; TAM 3) can be stated [13,14]. Thus, Peters et al [15] argued that this model alone does not indicate whether people would actually use a technology. In this context, basic human needs according to Ryan and Deci [16,17] and Deci and Ryan [18] play an important role. Following the Basic Psychological Need Theory [17], the more the interaction with the system satisfies basic psychological needs, the more the users will engage with a technology. Ryan and Deci [16,17] and Deci and Ryan [18], defined 3 basic psychological needs in their self-determination theory that are crucial to whether a person is proactive and engaged or passive and demotivated. These needs include competence (ie, feeling capable), relatedness (ie, feeling connected to others), and autonomy (ie, feeling self-determined). On the basis of this, Peters et al [15] defined the Motivation, Engagement, Thriving in User Experience (METUX) model, which can be used for the evaluation and iterative design process of technologies to increase motivation, engagement, and well-being. In this case, the 3 basic needs are mediating variables between a technical product and the well-being of the users. This implies that as autonomy increases, engagement increases; as competence increases, motivation increases; and as relatedness increases, well-being increases.

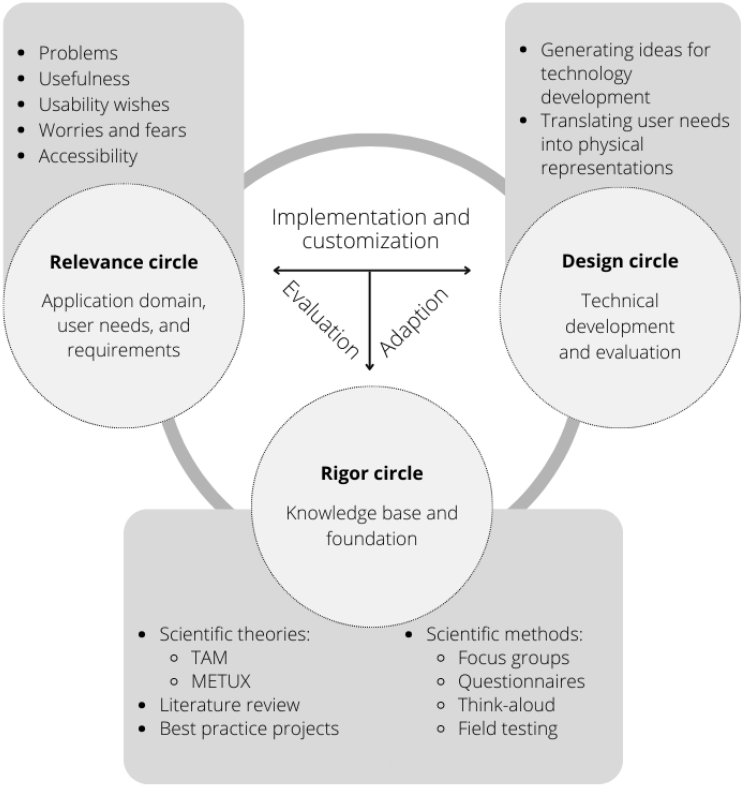
This Study

Acceptance building is a process that starts before the initial contact with the innovation and continues into the application phase, which was addressed within the pharmacy drone study (Apotheken-Drohnen-App; ADApp) [19]. This study represents a section of the whole ADApp project by investigating factors that are associated with the user acceptance of a drone-based medication delivery to be able to derive practice-relevant orientation points for participatory technology development (for apps and drones). Stephan et al [7] described that little attention is paid to the design phases of drones including the delivery process. Thus, this study used a mixed methods design and followed the methodological guidelines of the cocreative user-centered design proposed by Farao et al [20]. User-centered design is used to help consider the context of technologies as well as their implementation consequences at the design stage (Figure 1) [20,21]. Implementing technologies without user involvement may compromise the desired outcome, which in turn can lead to unmet health goals and adverse outcomes [20,22,23]. It is an evidence-based approach that involves users in developing technologies and prioritizes their needs [20,24]. In contrast to classical user-centered designs, this study used a controlled design for the first time. Traditionally, small sample groups are observed or interviewed or participate in usability tests during the testing and development phases of new technologies (usually operationalized through the think-aloud [TA] method and questionnaires). These are important approaches to get insights into key needs of the target population [25]. However, repeated measurements cause a change in the meaning of test scores, which makes it difficult to compare repeated measures. In a measurement perspective, it can be considered as bias in the measurement of change [26]. It can be inferred that conducting repeated measures with the same sample group may alter their attitudes, expectations, and behaviors in interacting with the technology. This could potentially influence their acceptance of the technology, as they become aware that the technology will adjust to their needs. Moreover, they are not unfamiliar with dealing with the app, which might influence user acceptance as well. This is indeed the purpose of a user-centered design, but it loses insights into perspectives of inexperienced users without a concrete expectation about the

XSL•FO
RenderX

changes in technology after giving feedback to it. By integrating a control group (called the *ad hocgroup*), this study aimed to investigate whether the assumption of the *core group* (ie, the same sample group at all measurement time points) is generalizable to a broader population. In this regard, a second purpose of this study was to answer the question of whether user acceptance differs between the core and ad hoc groups.

Figure 1. Implementation circle. METUX: Motivation, Engagement, Thriving in User Experience; TAM: technology acceptance model.



Methods

Study Design

The monocentric ADApp study aimed at an iterative and cocreative development of a pharmacy drone app with multiple measurement time points (Table 1). As this study design was embedded in the broader ADApp project, it was preceded by 2 research steps. As a first step, we conducted a scoping review of experimental studies examining the interaction between humans and drones during the delivery of drugs and defibrillators to identify research gaps and explore the scope of research activities [7]. In the next step, problems, needs, and requirements of the users were identified and concrete scenarios were outlined, which were necessary for the implementation of a first demonstrator of the app [27]. At this point, we decided to use focus groups instead of individual interviews because it allows participants to respond to each other's answers and gives us the most information. In this study, we conducted 3 iteration loops, where we tested the app demonstrators a well as the entire ADApp flow from order to delivery at an airport along with the user groups. The iterative process is one of the central features of the study and will be discussed in detail in the Study Setting section.

Table 1. The Apotheken-Drohnen-App design.

Goal	Methods	Participation	Time point	Reference
Knowledge	Scoping review	N/A[a]	August 2021	Stephan et al [7], 2022
Needs	Focus groups	User groups	October 6, 2021	Fink et al [27], 2023
Evaluation of functionalities	Questionnaires and think-aloud	First and second iteration loop: core group; 2 ad hoc groups	July 3, 2022	This study
Test flights	Questionnaire, think-aloud, and focus groups	Third iteration loop: core group; 1 ad hoc group	October 2022	This study

[a]N/A: not applicable.

To address potential biasing effects owing to a response shift through repeat interviews with the core group, a total of 3 ad hoc groups were acquired in this study. At each of the 3 development steps, a new and naive ad hoc group was used. This unique approach of adding ad hoc groups as control groups

XSL•FO
RenderX

also allows for the generalizability of key needs identified in the development process [25].

Study Setting

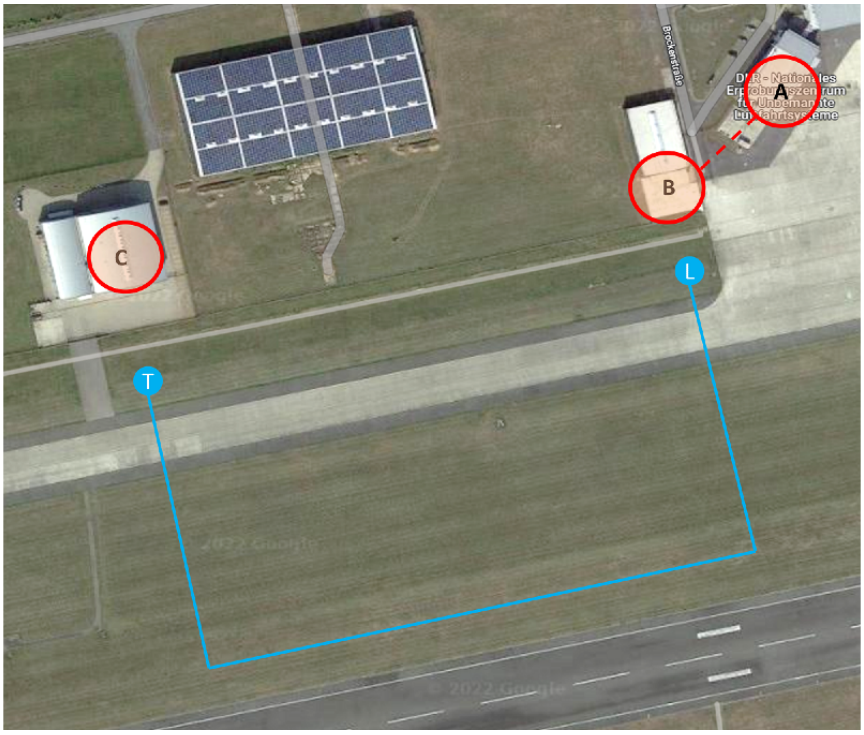
Our prior work showed that a pharmacy drone app, for example, must be lean and simple; must facilitate the user's performance (eg, software integration, shipment tracking, and plannability); must enable control (eg, handover identification and data sovereignty); and must include consultation and reconciliation features. The most frequently discussed problems associated with drones were the physical contact with the drone and the drone's noise [27]. Following the focus groups, the developers designed a demonstrator of the app including communication features (eg, an extra text field for communication with the pharmacy), plannability features (eg, time slots for drone delivery), shipment tracking, and features for enabling handover security, which was tested in 3 iteration loops with users. All loops were conducted separately per the core and ad hoc groups.

In the *first iteration loop* (conducted between March and April 2022), the usability of the app was reviewed. The app should be as easy as possible to use and allow users to create an account, add details such as delivery location, or submit an electronic prescription. The first iteration loop aimed at evaluating how intuitive the app is. For this purpose, users were shown the app and asked to think aloud while using it—without any introduction. For each participant, one experimenter prepared a protocol for taking notes. Finally, the prototypes were queried by using additional questionnaires: the Technology Usage Inventory (TUI) assessment was used to assess the intention to use, and the System Usability Scale (SUS) was used to measure usability. The basic psychological needs such as competence, autonomy, and connectedness were identified at the task level through the Technology-based Experience of Need Satisfaction (TENS) Task. After the first iteration loop, the demonstrator was adapted according to user feedback. The *second iteration loop* (conducted between June and July 2022) was dedicated to the design of the prototype and its technical development and evaluation. The decisive factor was how intuitive the design is and whether the tasks of the app are adequately represented. The procedure was the same as that in the first iteration loop, with the exception that instead of the TENS Task, basic psychological needs were queried via the TENS Interface. To gain deep insights into the participants' thoughts while using the technology, the core and ad hoc groups were again asked to think aloud while using the app. After receiving the second round of feedback from participants, a third demonstrator was developed. However, the *third iteration loop* (conducted in October 2022) aimed at testing the overall process starting with registration, setting the delivery location, submitting a prescription, and actually receiving a delivery. As focus groups showed concerns about injuries with the drone, this study wanted to test the handover via a winch, a parachute, or dropping to reduce physical contact with the drone, but unfortunately, owing to regulations and restrictions, testing the handover was not possible [27]. Thus, the drone had to land (Figure 2). Therefore, we divided participants into 2 groups: one group tested the app, whereas the other group talked about different handover scenarios and looked at the drone from close up. After each round, the groups were swapped. Similar to that in the first and second iteration loops, the intention to use and usability was queried using the TUI and SUS. To gauge the degree to which the technology fulfills users' needs in terms of behavior, participants completed the Basic Psychological Need Satisfaction and Frustration Scale (BPNSFS). Again, the participants were instructed to think aloud while going through the tasks in the app to gain more insights into the functionality of the app. After participants submitted their prescription, they went to the location in the airport where the drone was set to land (Figure 3). After landing, participants took the *medication* (a bag of gummy bears) out of the drone. After the testing, a short discussion was held with all participants to sum up their impressions.

Figure 2. Drone landing.

XSL•FO
RenderX

Figure 3. Airport landing zone. (A) Order location, (B) reception location, and (C) drone starting point. L: landing; T: takeoff.



Participants

Participants were chosen to represent potential user groups for both the drone-based medication delivery service and the supply chain, based on their respective role characteristics: physicians, nurses, pharmacists, and interested users, especially patients with COVID-19 and relatives of patients who need palliative care. Participants of the ad hoc group were recruited with the support of the Merseburger Innovations- und Technologiezentrum (innovation and technology center) and the ADApp project team. They were contacted via email or telephone. Core group participants were recruited from the focus group that was conducted in October 2021 [27]. A total of 3 pharmacists, 2 nurses, 3 general practitioners, and 1 patient with COVID-19 were recruited from the focus groups. Owing to the underrepresentation of interested users or patients, 2 more participants were recruited with the support of the ADApp project team.

All participants were informed about the general aim and reasons of the study and the procedure. They gave written informed consent before starting the iteration loops.

Ethics Approval

The ADApp study was approved by the Ethics Committee of the Martin-Luther University Halle-Wittenberg (protocol code 2021–069; date of approval May 6, 2021).

Measures

Quantitative Measures

SUS Score

The SUS measures the user's subjective perception of the usability of a system. It is technology independent, that is, it can be used for a wide range of systems and technologies [28-30]. Overall, 10 items were divided into 5 positively and 5 negatively directed statements, each represented on a 5-point Likert scale. The answers provided the SUS item score, which were then converted into the SUS overall score. The overall score ranges from 0 to 100. To calculate the SUS overall score, the first step was to calculate the raw score minus 1 for all odd items, whereas the raw score of 5 was subtracted for the even items. For example, if for item 1, the raw score is 4, the result is a score of 3 (4-1). If for item 2 the raw score is 2, the score is 3 (5-2). In the next step, the calculated scores are summed and multiplied by 2.5 [31].

Systems can be considered usable if they achieve the benchmark score of ≥68 [29,31]. In preliminary works, an adjective scale was developed for a more comprehensible classification of the SUS score, which ranges from outstanding (score 90-100) to very poor (score 0-34) [32].

TUI Assessment

The TUI assessment [33] aims to measure the intention to use and is based on the TAM [12]. The intention to use a technology is a comprehensive construct based on a variety of explanatory factors. As suggested in the TAM 2 and TAM 3, the TUI considers technology-specific factors and psychological factors. The TUI therefore supplements the classic technology acceptance factors of the TAM 1, such as perceived usability, usefulness, immersion, and accessibility (technology specific) with important psychological constructs, such as technology anxiety, curiosity, interest, and skepticism. The items, with the exception of the technology anxiety and interest scale, are related to a specific technology. Each scale consists of 3 to 4 items, each of which is to be rated on a 7-point Likert scale. The ninth scale measures the intention to use a specific technology with 3 items on a visual analog scale (each 100 mm). In total, the TUI consists of 33 items and has a modular design so that individual scales can be excluded and item formulations

XSL•FO
RenderX

can be adapted to the circumstances (eg, concrete technology names). With the exception of the "immersion" scale, all scales were used this study. All answers of a scale were summed to a sum value. It starts with 1 as the lowest expression of the construct and goes up to 21 (3 items) or 28 (4 items) as the highest expression, depending on the number of items. For the intention to use scale, the distance from the right end point (full rejection) to the answer cross on the line are measured. The distance in mm is determined for all 3 items and summed. The maximum scale value to be achieved is 300. The determined scale sum values are converted into standard values (stanines). The stanines reach from 1 (strongly below average) to 9 (strongly above average) [33].

METUX Model

The "pure" usability does not necessarily predict higher use of a technology [34]. Although the SUS questionnaire focuses on usability, the TUI questionnaire also includes psychological factors. However, both neglect basic psychological needs, which are addressed in the METUX model. It aims to optimize engagement, motivation, and well-being of technologies in iterative design processes [15]. Within the model, different spheres of experience were assumed to influence well-being: interface (ie, interacting with the technology), tasks (ie, engaging with the technology), behavior (ie, the relation to the overarching technology-supported behavior), and life (ie, the overall experience outside and beyond the technology). There are different questionnaires for measuring the basic psychological needs in different spheres. The spheres adoption, interface, tasks, and behavior were tested within this study using the Autonomy and Competence in Technology Adoption Questionnaire for the first adoption process, the TENS [15] for interface and task, and the BPNSFS [35,36] for behavior. The sphere "life" was not relevant for this study. The Autonomy and Competence in Technology Adoption Questionnaire addresses the question of why people use a technology and to what extent they experience themselves as competent to use it. It consists of 2 parts: the first, self-regulation, includes 12 items; and the second, perceived competence, includes 2 items, which are represented on a 5-point Likert scale. The goal of the TENS Interface and Task questionnaires is to assess the extent to which direct interaction (via interface) with a technology and engagement in technology-specific tasks satisfies the basic psychological needs for autonomy, competence, and relatedness [15]. In the TENS questionnaires, the items are each assigned to the basic needs of competence, autonomy, and relatedness but are presented randomly in the questionnaire. All items were equally weighted, summed, and averaged per basic need. The TENS Interface questionnaire consists of 15 items with 5 items per need, whereas the TENS Task questionnaire consists of 12 items with 4 items per need. The items are each represented on a 5-point Likert scale. As the TENS questionnaires are only validated for English-speaking countries so far, a linguistic validation of the questionnaires was conducted. For this purpose, the questionnaires were translated into German by an interpreter whose native language is German and who is fluent in English. Nevertheless, both the TENS questionnaires are not standardized for the German population that has to be considered when interpreting the results.

The BPNSFS is intended to assess the extent to which a technology improves need satisfaction in relation to the behavior the technology is intended to support [15]. The BPNSFS measures the satisfaction and frustration of the basic psychological needs of autonomy, competence, and relatedness. This includes a balanced combination of subscales for satisfaction and frustration. This distinction is necessary because the absence of need satisfaction does not equate to frustration of the same [17,35]. On the basis of the original questionnaire, Heissel et al [36] identified 6 different, but intercorrelating, factors with 4 items each for the German version of the BPNSFS: Autonomy Satisfaction, Autonomy Frustration, Competence Satisfaction, Competence Frustration, Relatedness Satisfaction, and Relatedness Frustration. Each response is represented on a 5-point Likert scale. The evaluation of the BPNSFS can be handled differently; for this study, the items per basic need are summed, and the 12 items of the subscales satisfaction and frustration are summed. There exist several adaptations of the BPNSFS, which have been validated and subjected to reliability testing. These adaptations include language translations, adjustment for age (children or adults), domain (sports, work, and romantic relationships), and clinical status (HIV, intellectual impairment, and chronic pain). However, a questionnaire related to technological aspects does not yet exist. Thus, for this study, the German version of the BPNSFS was used and minimally adapted according to the wording.

Qualitative Measures: TA Method

TA has traditionally been used in psychology and education to identify cognitive processes that occur internalized in the context of problem-solving [37]. In the context of technology development, TA is equally used to gain deep insights about thinking while using a technology [38]. The advantage of the method is to capture problems and solutions as the technology is being used, as retrospective surveys can lead to incomplete information about the problems of a technology. This means TA is helpful in tracing user thinking strategies [39,40]. For this purpose, participants were instructed to think aloud constantly while using the demonstrator. If participants stopped thinking aloud, they are reminded by the experimenter to continue speaking aloud [38,39]. For understanding problems and solutions, we did not explain how the demonstrator is supposed to work. Instead, we asked them to experience the app with little direction by explaining the task they had to do: registration, set delivery location, and submit recipe [41]. Thus, during the TA situation, it was important that the experimenter interact with participants as little as possible to prevent interference with the users' thoughts [39]. During the TA situations, the statements of the participants were digitally recorded (audio recordings) and transcribed afterward. Moreover, experimenters prepared protocols for making notes and describing events that are not verbally made by participants but important for analyzing. For example, if a participant said, "That is confusing," the experimenter protocoled what exactly was confusing [42].

XSL•FO
RenderX

Data Analysis

Quantitative Analyses

Analyses were performed with the statistical program SPSS Statistics (version 25; IBM Corp). Bivariate correlation analyses were performed to assess the association between different user acceptance measurements of usability (SUS score), intention to use a technology (TUI factors using raw values), and psychological needs (the TENS Task and Interface as well as the BPNSFS using raw values). To investigate which factors are associated with the intention to use the pharmacy drone app, hierarchical regression analyses were performed among TUI factors as well as among all measurements. To answer the question whether TUI and SUS factors differ between the core and ad hoc groups, 2-tailed *t* tests were performed. Owing to the nature of the study design (small sample sizes), statistical analyses should not be overinterpreted; thus, the results of questionnaires were also analyzed descriptively according to the improvement of acceptance level of SUS and TUI scores.

Qualitative Analyses

The transcripts were analyzed according to the event sampling method of Berelson [43], where an utterance represents an event. Utterances are defined as a complete sentence, a sentence fragment, or any sequence of speech separated in time (eg, a pause of 2 seconds) or semantic (eg, a change in content) [39,44]. Utterances were analyzed by referring phrase analysis [39]. First, all nouns and noun phrases were identified, and utterances with the reference concept name were coded by the first author (FF). This coding shows which concepts the participants focused on during the task. After concepts have been identified, these concepts were defined by the investigator (Tables 2 and 3). The resulting coding scheme was used to code the statements of the participants. Another researcher (JS) who was familiar with the analyses analyzed randomly selected portions of transcripts (20%) to determine if there was a match. In case of disagreement, discussions were held between the 2 examiners until a consensus was reached. Cohen κ [45] was computed for all variables. The interrater reliability was κ=0.654 (P<.001) with a substantial agreement [46].

Table 2. Examples of the coded concepts.

Coded concept	Segment
Value and problem	“What is stupid now is this field. It does not disappear.” (Doctor, aged 48 years)
Proposal	“Would be nicer: ‘Your location and address has been confirmed.’” (Nurse, aged 51 years)
Value and conceptuality	“That black sign that irritates.” (Patient, aged 55 years)
Status	“Now I have uploaded this successfully.” (Pharmacist, aged 35 years)
Ambiguity	“Do I have to register again now?” (Nurse, aged 49 years)

Table 3. Definitions of coded concepts.

Concept	Definition
Value	Rating of usefulness, importance, or worth
Status	Information indicative of status or self-instruction
Problem	Technical inconvenience requiring action
Ambiguity	Incomprehensibility of the process, operation, or handling
Conceptuality	System of terms or concepts
Proposal	Recommendation for technical implementation

The experimenters who prepared the protocols during the iteration loops were asked to evaluate the accuracy of the definitions of coded concepts to ensure that no undefined concepts remained. After all utterances were coded, concepts were summarized for groups. The results were arranged in a table per task and discussed with the ADApp team to derive practice-relevant orientation points. To rank the participant’s points, we classified the concepts according to criteria within the ADApp team. The criteria helped us to evaluate the relevance of the concepts for developing the technology. We defined 4 criteria: safety, risk in the delivery, optimization potential, and outside the capabilities (Multimedia Appendix 1).

Results

Participants

For the 3 iteration loops, between March 2021 and October 2022, we collected data from 18 participants (mean age 43.08, SD 12.44; range 25-65 years). Owing to the relatively large amount of time required, not all participants of the core group could always participate in the iteration loops. In total, 6 participants took part in the first core group (1 general practitioner, 2 pharmacists, 1 patient, and 2 nurses); 3 participants took part in the second core group (1 nurse, 1 pharmacist, and 1 patient); and 4 participants took part in the third core group (2 nurses, 1 pharmacist, and 1 patient). Although, we tried to balance the 2 groups, it was not always possible to equalize the core and ad hoc group. Four participants

XSL•FO
RenderX

took part in the first ad hoc group (1 general practitioner, 1 nurse, and 2 relatives of patients who need palliative care); 4 participants took part in the second ad hoc group (3 patients and 1 nurse); and 4 participants took part in the third ad hoc group (1 nurse and 3 patients). Table 4 provides detailed demographics.

Table 4. Participants' demographics.

Characteristics	Ad hoc group 1 (n[a]=4)	Ad hoc group 2 (n=4)	Ad hoc group 3 (n=4)	Core group 1 (n=6)	Core group 2 (n=3)	Core group 3 (n=4)	All (n=18)
Age (years), mean (SD; range)	50.25 (3.59; 47-55)	47.25 (4.42; 43-52)	50.75 (16.19; 31-64)	39.17 (15.45; 25-65)	37.00 (11.53; 26-49)	34.5 (10.97; 26-49)	46.00 (12.19; 25-65)
Gender (female)							
Values, mean (SD)	1.00 (0.00)	1.25 (0.50)	1.25 (0.50)	1.5 (0.55)	1.67 (0.57)	1.50 (0.58)	1.28 (0.46)
Values, n (%)	4 (100)	3 (75)	3 (75)	3 (50)	1 (33)	2 (50)	13 (72)
COVID-19							
Values, mean (SD)	1.5 (0.53)	1.25 (0.50)	1.50 (0.58)	1.5 (0.55)	1.00 (0.00)	1.25 (0.50)	1.44 (0.51)
Values, n (%)	2 (50)	3 (75)	2 (50)	3 (50)	3 (100)	3 (75)	10 (56)
Drone competence							
Values, mean (SD)	1.75 (0.50)	1.75 (0.50)	2.00 (0.00)	2.00 (0.00)	2.00 (0.00)	1.75 (0.50)	1.89 (0.32)
Values, n (%)	1 (25)	1 (25)	0 (0)	0 (0)	0 (0)	1 (25)	2 (11)
Medication app competence							
Values, mean (SD)	2.00 (0.00)	2.00 (0.00)	2.00 (0.00)	1.5 (0.55)	1.67 (0.57)	1.5 (0.58)	1.83 (0.38)
Values, n (%)	0 (0)	0 (0)	0 (0)	3 (50)	1 (33)	2 (50)	3 (17)

[a]n has been adjusted by weighting.

Quantitative Results

Bivariate Correlations

Bivariate correlations showed that the more usable (SUS) the technology was, the more the participants were willing to use it (TUI; r=0.833). Moreover, the more they found the technology useful (TUI), the more the technology was rated as usable (r=0.711; SUS and r=0.487; TUI) and the more participants would use the pharmacy drone app (TUI; r=0.754). In addition, the intention to use (TUI) the app was positively correlated with curiosity (TUI; r=0.550). The more skeptical (TUI) the participants were, the more participants rated the app as unusable (r=−0.542; SUS and r=−0.479; TUI) and would not be willing to use it (r=−0.446; TUI).

Furthermore, the intention to use (TUI) the technology was positively correlated with the perceived task competence (r=0.784; first iteration loop) and the need satisfaction of participant's competence (r=0.788; third iteration loop). The more participants felt competent in handling the app in the spheres of task (r=0.829; ie, engaging with the app; first iteration loop) and interface (r=0.929; ie, interacting with the app; second iteration loop), the higher was perceived usability (SUS). However, the more participants felt competent in task (r=0.675) and behavior (r=0.875), the more they felt autonomous (ie, the app provided options and participants did not feel under pressure from the app). The more curious the participants were, the more they felt satisfied in psychological needs of competence, autonomy, and relatedness (r=0.744). However, the absence of need satisfaction does not equate with the presence of need frustration [17,35]. The results revealed that the more participants rated the app as usable, the lesser they felt autonomy (r=−0.792; SUS and r=−0.768; TUI) and competence (r=−0.751; TUI) frustration. The more they believed the technology would be useful for them, the lesser they felt autonomy frustration (r=−0.827). Overall frustration shows a negative correlation with usability (r=−0.822; SUS and r=−0.799; TUI), usefulness (r=−0.923; TUI), and intention to use (r=−0.730; TUI). Moreover, the results indicated that autonomy frustration is a relevant marker for overall need frustration (r=0.933).

Furthermore, correlations showed that the more time participants needed to solve the task within the app, the less usable the app was rated (r=−0.534; SUS), the lesser they were willing to use the app (r=−0.429; TUI), the lesser they felt competent (r=−0.805), the more skepticism they had (r=0.504), and the older the participants were (r=0.681). The older the participants were, the more skeptical they were (r=.525) and the lesser they believed that the technology was accessible (r=−0.510). Female participants showed more technology anxiety (r=−0.505), had more interest (r=0.497), and felt more related by using the app (r=−0.800). The results of 2-tailed t tests confirmed these differences between female and male participants (anxiety: t_{19}=0.003; interest: $t_{23=}$0.012; relatedness satisfaction: t_6=0.004).

TUI Assessment

To test whether anxiety, curiosity, interest, skepticism, usefulness, usability, and accessibility contributed to higher intention to use, we regressed participant's ratings of these variables on their intention to use and controlled for age, gender, and duration using the app. Results showed that TUI factors such as usefulness, skepticism, and curiosity explain most of the variance in intention to use the pharmacy drone app

XSL•FO
RenderX

($F_{3,17}$=21.12; P<.001; R^2=0.788; adjusted R^2=0.751; Table 5). Usefulness (β=.499; P=.001) and curiosity (β=.376; P=.008) were significantly and positively associated with intention to use, whereas skepticism (β=−.397; P=.004) was significantly and negatively related with intention to use.

Table 5. Hierarchical regression analysis predicting intention to use the technology per the Technology Usage Inventory factors.

Model and predictor	Unstandardized coefficients, B (SE)	Standardized coefficients		R^2	R^2 change	F test (df)	P value
		β	P value				
1				0.582	0.560	26.48 (1,19)	<.001
Usefulness	14.73 (2.86)	.763	<.001				
2				0.676	0.640	18.81 (2,18)	<.001
Usefulness	13.03 (2.69)	.675	<.001				
Skepticism	−9.82 (4.29)	−.319	.03				
3				0.788	0.751	21.12 (3,17)	<.001
Usefulness	9.638 (2.51)	.499	.001				
Skepticism	−12.20 (3.66)	−.397	.004				
Curiosity	6.36 (2.12)	.376	.008				

Descriptively, anxiety, curiosity, interest, usefulness, and accessibility did not vary much between the core group and ad hoc groups, as shown in Table 6. Ad hoc groups appeared to be slightly more skeptical about the pharmacy drone app compared to the core group. This could potentially be attributed to age differences, as those in the ad hoc groups were older than those in the core group, and skepticism was positively correlated with age (Multimedia Appendix 2). Although usability remained at slightly below average in the ad hoc groups, it increased from average to slightly above average in the core group from the first to third iteration. Moreover, the core group showed higher ratings on the intention to use the pharmacy drone app than the ad hoc groups.

Table 6. Mean values of the Technology Usage Inventory stanine of the ad hoc and core groups per iteration loop.

	Ad hoc group 1	Ad hoc group 2	Ad hoc group 3	Core group 1	Core group 2	Core group 3
Psychological factors, mean (SD)[a]						
Anxiety	4.75 (1.89)	4.67 (3.05)	5.25 (1.50)	3.50 (1.22)	—[b]	3.00 (1.16)
Curiosity	7.00 (1.41)	7.67 (0.58)	5.25 (2.36)	7.33 (1.03)	—	7.00 (0.00)
Interest	5.50 (1.91)	5.25 (1.89)	7.25 (1.50)	6.00 (1.41)	5.67 (1.53)	6.00 (0.82)
Skepticism	3.75 (1.89)	2.75 (0.96)	3.50 (1.00)	2.83 (0.98)	1.67 (0.58)	2.00 (0.82)
Technology-specific factors, mean (SD)[a]						
Usefulness	8.00 (1.41)	8.75 (0.50)	8.00 (2.00)	8.67 (0.82)	9.00 (0.00)	8.75 (0.50)
Usability	4.50 (1.91)	4.25 (1.50)	4.50 (0.58)	5.50 (1.05)	6.67 (0.58)	7.00 (0.82)
Accessibility	7.75 (1.26)	4.50 (1.00)	6.25 (1.26)	7.33 (1.03)	8.67 (0.58)	7.25 (0.96)
Intention to use, mean (SD)[a]	6.75 (1.89)	8.50 (0.58)	5.75 (0.96)	8.50 (0.55)	8.00 (1.00)	8.75 (0.50)

[a]Stanine: 1 to 2, strongly below average; 3 to 4, slightly below average; 5, average; 6 to 7, slightly above average; and 8 to 9, strongly above average [33].

[b]Missing data.

The results of t tests showed a higher rating on usability for core group 3 compared with core group 1 (t_8=−2.68; P=.03). Moreover, independent samples t tests showed a higher anxiety (t_{19}=2.88; P=.01) as well as a higher skepticism toward the technology (t_{23}=2.17; P=.04) within the ad hoc group compared with the core group over all iteration loops. Moreover, the core group rated the app more usable than the ad hoc group (t_{23}=−3.33; P=.003) over all iteration loops. Although descriptive data showed a higher rating on intention to use the pharmacy drone app within the core group (mean 90.36, SD 11.01) than the ad hoc groups (mean 76.00, SD 23.74) over all iteration loops, 2-tailed t tests became not significant (t_3=−1.91; P=.07).

SUS Score

Descriptively, perceived usability (SUS score) decreased between ad hoc group 1 (rated as marginal) and ad hoc group 3 (rated as poor). Within the core group, the usability increased from core group 1 (rated as good) to core group 3 (rated as

XSL•FO
RenderX

excellent). However, within the second iteration loop, both groups (ad hoc and core groups) rated the app more usable than during the first and third iteration loops. Moreover, the core group showed higher SUS scores than the ad hoc groups (Table 7). Results of the *t* tests indicated a significant group difference in SUS score between ad hoc group 2 and ad hoc group 3 (t_6=3.35; *P*=.01). The results indicated lower SUS scores of ad hoc group 3 (rated as poor) than ad hoc group 2 (rated as excellent). Furthermore, independent samples *t* tests showed a significant group difference in the SUS score between the ad hoc and core groups over all iteration loops, with higher scores for the core group than the ad hoc groups (t_{23}=−2.87; *P*=.004).

Table 7. System Usability Score (SUS) scores (0 to 34: very poor; 35 to 49: poor; 50 to 67: marginal; 68 to 79: good; 80 to 89: excellent; and 90 to 100: outstanding [32]) in the ad hoc and core groups per iteration loop.

Group	SUS scores, mean (SD)
Ad hoc group 1	54.38 (23.22)
Ad hoc group 2	80.00 (14.43)
Ad hoc group 3	49.38 (11.25)
Core group 1	75.42 (16.84)
Core group 2	88.33 (10.10)
Core group 3	84.38 (5.54)
Ad hoc group 1 + core group 1	67.00 (21.34)
Ad hoc group 2 + core group 2	83.57 (11.57)
Ad hoc group 1 + core group 3	66.88 (20.43)
All groups	71.60 (19.75)

Overall Regressions

When regressing all factors of all usability and psychological needs, the results showed that the SUS usability score as well as the TUI factors such as curiosity and interest explained most of the variance on intention to use the pharmacy drone app ($F_{3,16}$=40.27; *P*<.001; R^2=.883; adjusted R^2=.861; Table 8). Usability (β=.845; *P*<.001), curiosity (β=.232; *P*=.02), and interest (β=.195; *P*=.04) were significantly and positively associated with intention to use.

Table 8. Hierarchical regression analysis predicting intention to use the technology per the Technology Usage Inventory, System Usability Scale (SUS), Technology-based Experience of Need Satisfaction Task, and Basic Psychological Need Satisfaction and Frustration Scale factors.

Model and predictor		Unstandardized coefficients, B (SE)	Standardized coefficients		R^2	R^2 change	*F* test (*df*)	*P* value
			β	*P* value				
1					0.789	0.777	67.16 (1,18)	<.001
	SUS usability	0.999 (0.122)	.888	<.001				
2					0.848	0.830	47.31 (2,17)	<.001
	SUS usability	0.890 (0.115)	.791	<.001				
	Curiosity	4.45 (1.73)	.262	.02				
3					0.883	0.861	40.27 (3,16)	<.001
	SUS usability	0.951 (0.107)	.845	<.001				
	Curiosity	3.94 (1.58)	.232	.02				
	Interest	2.88 (1.31)	.195	.04				

Qualitative Results

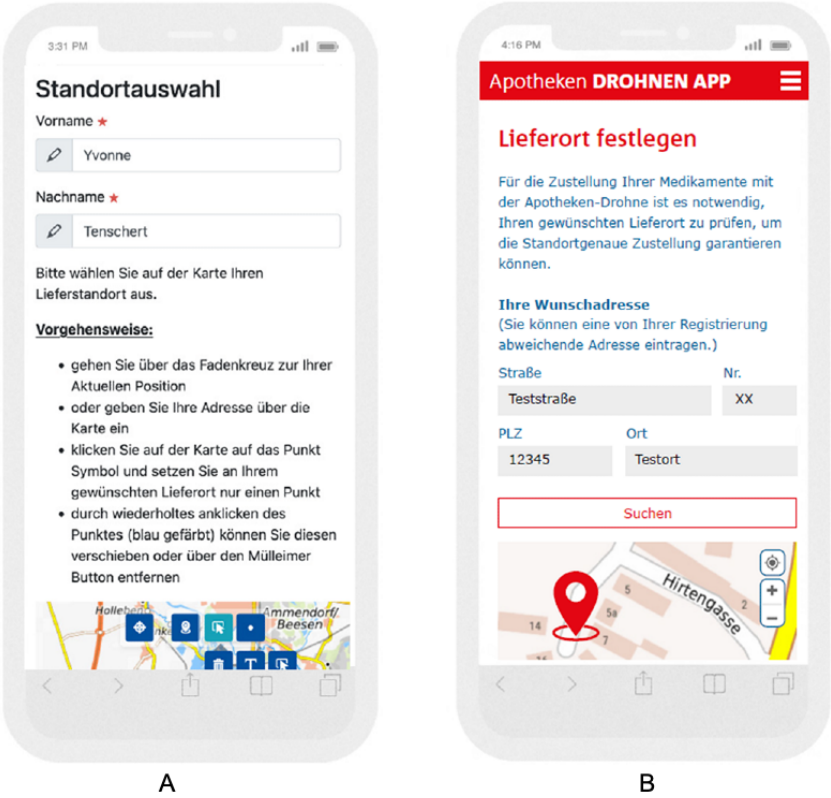
First Iteration

After the first iteration, the app received a new design according to user's feedback. Important points after the first iteration were providing more guidance through the app with information about next steps and reasons for doing these steps (eg, the importance of setting the delivery location, clear information about how to choose the delivery location, and more precise symbols; Figures 4 and 5); adaption of conceptualizations (eg, "location" [German: *Standort*] to "delivery point" [German: *Lieferort*]; Figure 4); automatizations (eg, transferring address data automatically to the map); minimizations (eg, reducing symbols within the map and information within each step); communication options (eg, integrating a field to formulate a message to the pharmacist); control features (eg, an order summary); and autonomy options (eg, to upload >1 prescription if necessary). However, participants missed a visualization of password requirements and a preview function of the uploaded recipe. Furthermore, users would rather preview individual

XSL•FO
RenderX

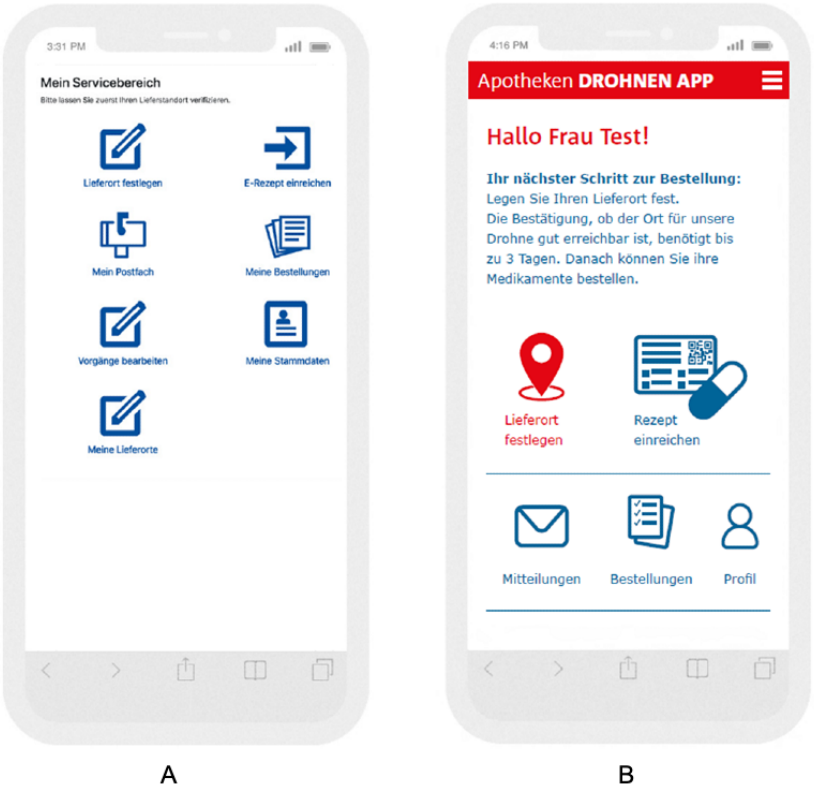
pages per click than perform 1-page scrolling. They would also desire information about payment options within the app as well as details about medication availability. This necessitates integration of the interface with the pharmacy's merchandise management system, which requires additional technical and regulatory administration. A short solution for that was to integrate a comment field at the step of ordering the medication to describe further medication wishes as well as to ask questions to the pharmacist. Moreover, participants emphasized, at this point, the importance of shipment tracking, as Fink et al [27] described in their study. However, the most difficult step participants reported was setting of the delivery location. This was also shown in the amount of support needed while using the app (Multimedia Appendix 3). Although experimenters were instructed to not help participants, at some points the help was necessary so that the participants could finish the task.

Figure 4. User-centered app design (A) before and (B) after first iteration: delivery location.



XSL•FO
RenderX

Figure 5. User-centered app design (A) before and (B) after first iteration: symbols.



Second Iteration

After the second iteration, participants in both the core group and the ad hoc group reported that registration was simplified, indicating that it was easy for them to register. Moreover, participants reported that the summary of order was clear, the texts were more comprehensible (core group), and the participants liked the option to upload >1 prescription (ad hoc group). However, participants missed a preview function of the uploaded prescription as a picture and the selection of push messages. Moreover, they suggested highlighting the icon, which is the next step to be clicked on (eg, after setting the delivery location, the icon for submitting the prescription should be highlighted), and highlighting the inbox when a new message has arrived. Important points after the second iteration were still reducing symbols within the map (eg, symbols were reduced to a minimum, and the text describing the symbols was shortened because the participants did not read the instructions above the card); more guidance through the app with information about next steps; and more transparency about how a response will be received from pharmacists (eg, integration of information about how the contact will take place, via mail or telephone), indications on password requirements (such as length, upper and lower case, and special characters), and information about shipment tracking. An important safety-relevant point was that flight slots were not up to date according to the original time. Similar to the first iteration loop, participants reported setting the delivery location as the most difficult step. They would like to have more guidance for the subsequent steps after determination of the delivery location.

Third Iteration

The core group reported that the app was more intuitive and had improved compared with the first iteration loop. They mentioned that the automated fill-in of address data in the map as well as the identification that the prescription was successfully uploaded was useful.

Important points after the third iteration were more guidance through a processing status within the app; adaption of conceptualization (eg, “mailbox” [German: *Postfach*] to “messages” [German: *Nachrichten*]); visualization of password requirements (eg, upper and lower case and special characters); automatization features (eg, automatic suggestions such as city when entering the postal code); differentiation (eg, distinction between the delivery and billing address); control (eg, adjustment of the amount of information within the app to control how much information the user wants, which might be configured via profile); push messages instead of emails because participants did not read the texts despite shortening them; a preview function of the uploaded prescription; and design features (eg, it was not clear that scrolling was necessary, thus more guidance is useful through individual pages [click to continue]). They also desired the inbox to be highlighted when new messages had arrived.

Although the app was adapted according to the participant’s feedback, the ad hoc group reported that the step of setting the delivery location remained too complex for them. They mentioned that this step was too bulky, time consuming, and not intuitive despite adjustments such as the reduction of map symbols, providing the most important information, shortening the text, and inserting address data automatically in the map. Participants of the ad hoc group felt lost and helpless during

XSL•FO
RenderX

this step. However, during discussions after thinking aloud, participants suggested that the setting of delivery comes from the operator and thus must not be made by users. They only wanted confirmation of the delivery point to ensure its correctness. Excluding this step might decrease the likelihood of user errors.

However, across all iterations, criteria emerged that were repeatedly the focus of the participants' attention: automatization (ie, easy and fast use for avoiding redundancies), minimization (ie, as little information as possible and as much as necessary), differentiation (ie, clear distinctions), control (ie, options to choose), guiding (ie, concise and understandable instructions supporting guidance through the app), conceptualization (ie, easy and precise language), barrier-free design (ie, uniformity between different steps and intuitive visualizations), and transparency (about disclosures to be made or obtained information).

Although we could not test the handover scenarios, we have made modifications to the drone. Previous results of focus group testing within the ADApp project showed concerns about injuries caused by the drone [27]. Thus, the drone has now been given a flap underneath so that the medications can be dropped by ejection, using a parachute, or using a winch and a landing of the drone can be prevented (Figure 6). Further testing is planned to test different handover scenarios with participants to adapt the handover according to their needs.

Figure 6. Drone medication ejection.

Discussion

Factors of User Acceptance

This study aimed at investigating factors that are associated with the user acceptance of a drone-based medication delivery by using a mixed methods design to be able to derive practice-relevant orientation points for participatory technology development (for apps and drones).

First, an important point is duration handling with the app. Older participants needed more time to solve the tasks within the app. Furthermore, the longer the process took, the more the usability, intention to use the app, and feelings of competence decreased, while skepticism increased. Therefore, the duration of interaction with a technology appears to be a crucial factor for user acceptance.

Second, psychological factors such as skepticism and curiosity as well as technology-specific factors such as usefulness and usability are related with participants' intention to use the technology for a drone-based medication delivery. Regression analyses within the TUI factors revealed usefulness, curiosity, and skepticism as significant predictors for intention to use the technology, wherein usefulness explained the highest variance (49.9%), which is consistent with the findings of Güsken et al [47]. This implies a particular relevance of factor usefulness for the development of technologies in health care, especially in drone-based medication delivery. With the help of the TUI questionnaire, we can conclude that curiosity and skepticism affect user acceptance. The more the users were curious about the pharmacy drone app and the less skeptical they were, the more the users were willing to use and interact with the technology. This is consistent with the findings of Eißfeld et al [6] who found that a positive attitude toward drones and a general technical interest are related to improved information about it.

Third, basic human needs according to Ryan and Deci [16,17] and Deci and Ryan [18]also play an important role. Results of this study showed that the more participants perceived competence in handling the app, the more they are willing to use the technology and the more they rated the app as usable. This implies that, although competence satisfaction in all iteration loops was related with usability and the intention to use, autonomy and relatedness were not related. Nevertheless, results showed that the more participants felt competent, the more they felt autonomous. Moreover, the basic psychological needs (competence, autonomy, and relatedness) were positively correlated with curiosity. In addition, the lesser participants felt frustrated on psychological needs, the higher they rated the usability, usefulness, and intention to use. Thus, following the Basic Psychological Need Theory [17], the more the interaction with the system satisfies basic psychological needs, the more the users will engage with a technology. Following the METUX model, an increase in autonomy increases engagement and an increase in competence increases motivation of using the app, which is in accordance with the results of this study [15]. Interestingly, while competence and autonomy appear to be significant factors in explaining differences in intention to use and usability, relatedness does not play a role, despite focus

XSL•FO
RenderX

group discussions emphasizing the importance of communication and consultation features within such technology [27]. One reason for this result might be the nature of the study task. Although participants used the app in a simulated context, communication aspects did not play a role in this developing step of technology. Therefore, supply studies in real-life contexts are necessary to test the impact of psychological and technological factors on real-life complex problems, which cannot be fully investigated in a simulated context such as in this study, where, for example, relatedness might hold greater significance in real-life scenarios than in simulations, as increased relatedness could potentially enhance overall well-being [15].

Fourth, overall regression analyses showed that usability, curiosity, and interest explain most of the variance on intention to use the pharmacy drone app, wherein usability showed the strongest effect (84.5%). This means, when adding all factors in one model, usability becomes the strongest predictor for intention to use the pharmacy drone app. Similar to other studies, this study found evidence for the importance of usability in using a technology [1,5]. This means that better usability of a technology leads to higher acceptance [12,33]. However, this study used a very small sample group, and the statistical results should be considered carefully. In conclusion, studies with higher sample sizes are necessary.

Taken together, the intention to use a drone-based medication delivery system is a comprehensive construct that is based on a large number of underlying, explanatory factors. This study showed that usability, curiosity, and interest had a considerable impact on intention to use, wherein usefulness, skepticism, and competence also played an important role. The failure to actively involve users in technology development can thus result in insufficient addressing of profession- and person-specific needs, thus resulting in a lack of intention to use the technology. For successful technology development, it is therefore crucial to develop an understanding of the necessary characteristics of health care technologies and to identify the determinants that ensure a high level of acceptance for improving the current supply situation [47].

In accordance with the previous scoping review [7], user feedback was collected iteratively and focused on user experience. The TA method [38] used in this process provided valuable insights that were taken into account when developing the app. In this way, important changes to the app were successfully implemented by user request such as a reduced design and automatic fill-in aids. It was found that communication with the dispatcher and shipment tracking are very important to users, which is consistent with the assumptions from the scoping review [7], and this also led to further adjustments. The changes made could be verified in the further iterations; for example, it turned out that the revision for the definition of the delivery location was not helpful: the process was adjusted based on participant feedback with more information, but this step remained too complex. In the third iteration, it became clear that the required texts were not being read at this point, leading to the ultimate decision to omit this step altogether as participants indicated that they only wanted to confirm the delivery location.

Differences in User Acceptance

A second purpose of this study was to assess group differences between a core group and ad hoc groups in user acceptance.

The results of this study indicated the importance of an ad hoc group in an iterative, cocreative process. Although within the core group, intention to use (TUI) was similar over all iteration loops (strongly above average), within the ad hoc group, intention to use varied from slightly above average to strongly above average to average. Although within the core group, the usability of the app slightly increased from “average resp. good” to “slightly above average resp. excellent,” the usability within the ad hoc group decreased from marginal to poor (SUS) and remained slightly above average (TUI). Results of *t* tests of both questionnaires showed a significant group difference between the ad hoc and core groups, with higher ratings in perceived usability for the core group. This suggests that repeated measurements induce a shift in the interpretation of test scores, potentially biasing the measurement of change [26]. Thus, this study shows that repeated measures with the same sample group might change their attitude, expectations, and behavior in dealing with the technology, which changed their ratings on usability. The core group then tended to evaluate the app better than the ad hoc group because they were not unfamiliar with dealing with the app. Thus, for naive users, the app is just not intuitive and easy enough to use. An additional explanation for this result is that the ad hoc groups were more anxious and skeptical than the core group, wherein a higher skepticism was found to be related to lower ratings on usability [47]. However, within the usability score (SUS), data showed an increase from the first to second iteration and a decrease from the second to third iteration in both groups, wherein the ad hoc groups rated the pharmacy drone app as less usable compared with the core group. One reason for the decrease from the second to third iteration could be the more complex setting during the third iteration: the participants had to run through the entire process from registration, setting the delivery location, and ordering the medication to receiving the medication. Meanwhile, the core group showed a learning effect and maybe thus rated the pharmacy drone process to be more usable from the first iteration to second iteration, and the ad hoc groups could stumble because of the complexity.

Limitations

This study shows for the first time the importance of ad hoc groups as a control group while developing and evaluating a technology in a user-centered design. When interpreting the results of this study, several methodological limitations must be considered. In terms of age and gender distribution, the sample can be classified as unrepresentative owing to the small number of participants. This is particularly evident in the statistical evaluation. Nevertheless, there is a basic tendency toward a clear effect, which is evident despite the small number of samples. The participants had a basic interest in new topics and in the topic itself. Although the risk of “positive selection” cannot be completely ruled out, it is not seen because the topic of drone-assisted medication delivery was largely unknown. Thus, the perspectives of participants who consistently reject technical systems in the context of care and delivery were as

XSL•FO
RenderX

poorly represented as those who chose not to participate in the surveys for other reasons. Reasons for this could include a general dismissive attitude toward additional effort owing to time resources, heavy workloads, or other thematic priorities.

However, it can be concluded that the results obtained to assess the acceptance of the drone app for utility purposes have revealed important insights regarding technical development and its practical use. In this context, the findings exhibit similarities to surveys conducted for other target groups in health care.

Conclusions

The study highlights the significance of understanding the essential attributes of health care technologies and the factors that lead to their acceptance in improving the current supply situation. It offers valuable insights for practitioners to develop participatory technologies and recommends ad hoc groups as a complementary approach to control the process of a user-centered design. However, larger samples and real-world contexts are required to confirm these findings.

Acknowledgments

The authors would like to thank the Apotheken-Drohnen-App team for their involvement in the participatory development of the drone. In addition, the authors would like to thank Dr Anne-Marie Lachmund for her help with proofreading.

Authors' Contributions

FF contributed to conceptualization, formal analysis, investigation, methodology, visualization, and writing of the original draft. IK contributed to visualization and writing of the original draft. JVS helped with the formal analysis. PJ contributed to conceptualization, funding acquisition, supervision, and writing of the original draft (review and editing). HKH helped with the formal analysis and visualization. DP contributed with writing of the original draft (review and editing).

Conflicts of Interest

None declared.

Multimedia Appendix 1
Examples of criteria.
[DOCX File , 13 KB - humanfactors_v11i1e51587_app1.docx]

Multimedia Appendix 2
Bivariate correlations between System Usability Scale (SUS), Technology Usage Inventory (TUI), Technology-based Experience of Need Satisfaction (TENS) Task, TENS Interface, and Basic Psychological Need Satisfaction and Frustration Scale (BPNSFS).
[DOCX File , 46 KB - humanfactors_v11i1e51587_app2.docx]

Multimedia Appendix 3
Support required per task.
[DOCX File , 13 KB - humanfactors_v11i1e51587_app3.docx]

References

1. Hiebert B, Nouvet E, Jeyabalan V, Donelle L. The application of drones in healthcare and health-related services in North America: a scoping review. Drones 2020 Jul 04;4(3):30. [doi: 10.3390/drones4030030]
2. Saponi M, Borboni A, Adamini R, Faglia R, Amici C. Embedded payload solutions in UAVs for medium and small package delivery. Machines 2022 Aug 27;10(9):737. [doi: 10.3390/machines10090737]
3. Sabino H, Almeida RV, Moraes LB, Silva WP, Guerra R, Malcher C, et al. A systematic literature review on the main factors for public acceptance of drones. Technol Soc 2022 Nov;71:102097. [doi: 10.1016/j.techsoc.2022.102097]
4. Rave A, Fontaine P, Kuhn H. Drone network design for emergency resupply of pharmacies and ambulances. SSRN. Preprint posted online September 14, 2023 [FREE Full text] [doi: 10.2139/ssrn.4569199]
5. Rejeb A, Rejeb K, Simske S, Treiblmaier H. Humanitarian drones: a review and research agenda. Internet Things 2021 Dec;16:100434. [doi: 10.1016/j.iot.2021.100434]
6. Eißfeldt H, Vogelpohl V, Stolz M, Papenfuß A, Biella M, Belz J, et al. The acceptance of civil drones in Germany. CEAS Aeronaut J 2020 Apr 04;11(3):665-676. [doi: 10.1007/s13272-020-00447-w]
7. Stephan F, Reinsperger N, Grünthal M, Paulicke D, Jahn P. Human drone interaction in delivery of medical supplies: a scoping review of experimental studies. PLoS One 2022;17(4):e0267664 [FREE Full text] [doi: 10.1371/journal.pone.0267664] [Medline: 35482656]

XSL•FO
RenderX

8. Sham R, Siau CS, Tan S, Kiu DC, Sabhi H, Thew HZ, et al. Drone usage for medicine and vaccine delivery during the COVID-19 pandemic: attitude of health care workers in rural medical centres. Drones 2022 Apr 27;6(5):109 [FREE Full text] [doi: 10.3390/drones6050109]
9. Milchrahm E. Entwicklung eines Modells zur Akzeptanzproblematik von Informationstechnologie. In: Proceedings of the Information und Mobilität, Optimierung und Vermeidung von Mobilität durch Information. 2002 Presented at: ISI 2002; October 8-10, 2002; Regensburg, Germany.
10. Krey M, Seiler R. Usage and acceptance of drone technology in healthcare: exploring patients and physicians perspective. In: Proceedings of the 52nd Hawaii International Conference on System Sciences. 2019 Presented at: HICSS 2020; January 8-11, 2019; Grand Wailea, HI URL: https://digitalcollection.zhaw.ch/handle/11475/15635
11. Simon B. E-Learning an Hochschulen. Gestaltungsräume und Erfolgsfaktoren von Wissensmedien. Lohmar, Germany: Josef Eul Verlag; 2001.
12. Davis FD. Perceived usefulness, perceived ease of use, and user acceptance of information technology. MIS Q 1989 Sep;13(3):319-340. [doi: 10.2307/249008]
13. Venkatesh V, Bala H. Technology Acceptance Model 3 and a research agenda on interventions. Decis Sci 2008 May 09;39(2):273-315. [doi: 10.1111/j.1540-5915.2008.00192.x]
14. Venkatesh V, Davis FD. A theoretical extension of the technology acceptance model: four longitudinal field studies. Manag Sci 2000 Feb;46(2):186-204. [doi: 10.1287/mnsc.46.2.186.11926]
15. Peters D, Calvo RA, Ryan RM. Designing for motivation, engagement and wellbeing in digital experience. Front Psychol 2018;9:797 [FREE Full text] [doi: 10.3389/fpsyg.2018.00797] [Medline: 29892246]
16. Ryan RM, Deci EL. Intrinsic and extrinsic motivations: classic definitions and new directions. Contemp Educ Psychol 2000 Jan;25(1):54-67. [doi: 10.1006/ceps.1999.1020] [Medline: 10620381]
17. Ryan RM, Deci EL. Self-Determination Theory: Basic Psychological Needs in Motivation, Development, and Wellness. New York, NY: Guilford Press; 2017.
18. Deci EL, Ryan RM. Intrinsic Motivation and Self-Determination in Human Behavior. New York, NY: Springer; 1985.
19. Jockisch M. Das ist gar kein Modell!. In: Bandow G, Holzmüller H, editors. Das Technologieakzeptanzmodell. Wiesbaden, Germany: Springer Gabler; 2010:233-254.
20. Farao J, Malila B, Conrad N, Mutsvangwa T, Rangaka MX, Douglas TS. A user-centred design framework for mHealth. PLoS One 2020;15(8):e0237910 [FREE Full text] [doi: 10.1371/journal.pone.0237910] [Medline: 32813711]
21. Risling TL, Risling DE. Advancing nursing participation in user-centred design. J Res Nurs 2020 May 03;25(3):226-238 [FREE Full text] [doi: 10.1177/1744987120913590] [Medline: 34394630]
22. Schnall R, Rojas M, Bakken S, Brown W, Carballo-Dieguez A, Carry M, et al. A user-centered model for designing consumer mobile health (mHealth) applications (apps). J Biomed Inform 2016 Apr;60:243-251 [FREE Full text] [doi: 10.1016/j.jbi.2016.02.002] [Medline: 26903153]
23. Nilsen W, Kumar S, Shar A, Varoquiers C, Wiley T, Riley WT, et al. Advancing the science of mHealth. J Health Commun 2012;17 Suppl 1:5-10. [doi: 10.1080/10810730.2012.677394] [Medline: 22548593]
24. McCurdie T, Taneva S, Casselman M, Yeung M, McDaniel C, Ho W, et al. mHealth consumer apps: the case for user-centered design. Biomed Instrum Technol 2012;Suppl:49-56. [doi: 10.2345/0899-8205-46.s2.49] [Medline: 23039777]
25. Altman M, Huang TT, Breland JY. Design thinking in health care. Prev Chronic Dis 2018 Sep 27;15:E117 [FREE Full text] [doi: 10.5888/pcd15.180128] [Medline: 30264690]
26. Oort FJ, Visser MR, Sprangers MA. Formal definitions of measurement bias and explanation bias clarify measurement and conceptual perspectives on response shift. J Clin Epidemiol 2009 Nov;62(11):1126-1137. [doi: 10.1016/j.jclinepi.2009.03.013] [Medline: 19540722]
27. Fink F, Paulicke D, Grünthal M, Jahn P. "Of course, drones delivering urgent medicines are necessary. But I would not use them until…" Insights from a qualitative study on users' needs and requirements regarding the use of medical drones. PLoS One 2023;18(5):e0285393 [FREE Full text] [doi: 10.1371/journal.pone.0285393] [Medline: 37155650]
28. Brooke J. SUS: a 'quick and dirty' usability scale. In: Usability Evaluation In Industry. Boca Raton, FL: CRC Press; 1996.
29. Brooke J. SUS: a retrospective. J Usab Stud 2013 Feb;8(2):29-40.
30. Hyzy M, Bond R, Mulvenna M, Bai L, Dix A, Leigh S, et al. System usability scale benchmarking for digital health apps: meta-analysis. JMIR Mhealth Uhealth 2022 Aug 18;10(8):e37290 [FREE Full text] [doi: 10.2196/37290] [Medline: 35980732]
31. Sauro J. A Practical Guide to the System Usability Scale: Background, Benchmarks & Best Practices. Denver, CO: Measuring Usability LLC; 2011.
32. Bangor A, Kortum P, Miller J. Determining what individual SUS scores mean: adding an adjective rating scale. J Usab Stud 2009 May 1;4(3):114-123.
33. Kothgassner OD, Felnhofer A, Hauk N, Kastenhofer E, Gomm J, Ryspin-Exner I. TUI: Technology Usage Inventory. Vienna, Austria: ICARUS; 2012.
34. Febretti A, Garzotto F. Usability, playability, and long-term engagement in computer games. In: Proceedings of the CHI '09 Extended Abstracts on Human Factors in Computing Systems. 2009 Presented at: CHI EA '09; April 4-9, 2009; Boston, MA. [doi: 10.1145/1520340.1520618]

XSL•FO
RenderX

35. Chen B, Vansteenkiste M, Beyers W, Boone L, Deci EL, Van der Kaap-Deeder J, et al. Basic psychological need satisfaction, need frustration, and need strength across four cultures. Motiv Emot 2015;39(2):216-236. [doi: 10.1007/s11031-014-9450-1]
36. Heissel A, Pietrek A, Flunger B, Fydrich T, Rapp MA, Heinzel S, et al. The validation of the German basic psychological need satisfaction and frustration scale in the context of mental health. Eur J Health Psychol 2018 Oct;25(4):119-132. [doi: 10.1027/2512-8442/A000017]
37. Gunzenhauser C, Karbach J, Saalbach H. Function of verbal strategies in monolingual vs. bilingual students' planning performance: an experimental approach. Cognit Dev 2019 Apr;50:1-12. [doi: 10.1016/j.cogdev.2019.01.003]
38. Jaspers MW, Steen T, van den Bos C, Geenen M. The think aloud method: a guide to user interface design. Int J Med Inform 2004 Nov;73(11-12):781-795. [doi: 10.1016/j.ijmedinf.2004.08.003] [Medline: 15491929]
39. Fonteyn ME, Kuipers B, Grobe SJ. A description of think aloud method and protocol analysis. Qual Health Res 2016 Jul 01;3(4):430-441. [doi: 10.1177/104973239300300403]
40. Johnson WR, Artino ARJ, Durning SJ. Using the think aloud protocol in health professions education: an interview method for exploring thought processes: AMEE Guide No. 151. Med Teach 2022 Dec 19;45(9):937-948. [doi: 10.1080/0142159x.2022.2155123]
41. Roberts JP, Fisher TR, Trowbridge MJ, Bent C. A design thinking framework for healthcare management and innovation. Healthc (Amst) 2016 Mar;4(1):11-14. [doi: 10.1016/j.hjdsi.2015.12.002] [Medline: 27001093]
42. Ericsson KA. Protocol analysis and expert thought: concurrent verbalizations of thinking during experts' performance on representative tasks. In: Ericsson KA, Charness N, Feltovich PJ, Hoffman RR, editors. The Cambridge Handbook of Expertise and Expert Performance. Cambridge, MA: Cambridge University Press; 2006:223-242.
43. Berelson B. Content Analysis in Communication Research. Glencoe, Ill: Free Press; 1952.
44. Winsler A, Fernyhough C, McClaren EM, Way E. Private speech coding manual. Unpublished manuscript. George Mason University. 2005. URL: https://www.researchgate.net/profile/Adam-Winsler/publication/228654613_Private_speech_coding_manual/links/09e4150d729fe68323000000/Private-speech-coding-manual.pdf [accessed 2024-03-29]
45. Cohen J. A coefficient of agreement for nominal scales. Educ Psychol Measure 1960;20(1):37-46. [doi: 10.1177/001316446002000104]
46. Landis JR, Koch GG. The measurement of observer agreement for categorical data. Biometrics 1977 Mar;33(1):159-174. [Medline: 843571]
47. Güsken SR, Frings K, Zafar F, Saltan T, Fuchs-Frohnhofen P, Bitter-Krahe J. Einflussfaktoren auf die Nutzungsintention von Pflegekräften zur Verwendung digitaler Technologien in der ambulanten Pflege. Fallstudie zur Einführung eines Sensortexils. Z Arbeitswiss 2021 Oct 13;75(4):470-490 [FREE Full text] [doi: 10.1007/s41449-021-00277-4] [Medline: 34658493]

Abbreviations

ADApp: Apotheken-Drohnen-App
BPNSFS: Basic Psychological Need Satisfaction and Frustration Scale
METUX: Motivation, Engagement, Thriving in User Experience
SUS: System Usability Scale
TAM: technology acceptance model
TA: think-aloud
TENS: Technology-based Experience of Need Satisfaction
TUI: Technology Usage Inventory

Edited by A Kushniruk; submitted 04.08.23; peer-reviewed by S Ariffin, C Baxter, J Brooke; comments to author 29.10.23; revised version received 21.12.23; accepted 06.01.24; published 30.04.24.

Please cite as:
Fink F, Kalter I, Steindorff JV, Helmbold HK, Paulicke D, Jahn P
Identifying Factors of User Acceptance of a Drone-Based Medication Delivery: User-Centered Design Approach
JMIR Hum Factors 2024;11:e51587
URL: https://humanfactors.jmir.org/2024/1/e51587
doi:10.2196/51587
PMID:38687589

©Franziska Fink, Ivonne Kalter, Jenny-Victoria Steindorff, Hans Konrad Helmbold, Denny Paulicke, Patrick Jahn. Originally published in JMIR Human Factors (https://humanfactors.jmir.org), 30.04.2024. This is an open-access article distributed under

XSL•FO
RenderX

the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on https://humanfactors.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Facilitators of and Barriers to the Use of a Digital Self-Management Service for Diagnostic Testing: Focus Group Study With Potential Users

Kyma Schnoor[1,2], MSc; Esther P W A Talboom-Kamp[1,2,3], MBA, MD, PhD; Muamer Hajtić[4], MSc; Niels H Chavannes[1,2], MD, PhD; Anke Versluis[1,2], PhD

[1]Department of Public Health and Primary Care, Leiden University Medical Center, Leiden, Netherlands
[2]National eHealth Living Lab, Leiden University Medical Center, Leiden, Netherlands
[3]Zuyderland, Sittard-Geleen, Netherlands
[4]Department of Medical informatics, University of Amsterdam, Amsterdam, Netherlands

Corresponding Author:
Kyma Schnoor, MSc
Department of Public Health and Primary Care
Leiden University Medical Center
Hippocratespad 21
Leiden
Netherlands
Phone: 31 71526 8433
Email: k.schnoor@lumc.nl

Abstract

Background: Health care lags in digital transformation, despite the potential of technology to improve the well-being of individuals. The COVID-19 pandemic has accelerated the uptake of technology in health care and increased individuals' willingness to perform self-management using technology. A web-based service, Directlab Online, provides consumers with direct digital access to diagnostic test packages, which can digitally support the self-management of health.

Objective: This study aims to identify the facilitators, barriers, and needs of Directlab Online, a self-management service for web-based access to diagnostic testing.

Methods: A qualitative method was used from a potential user's perspective. The needs and future needs for, facilitators of, and barriers to the use of Directlab Online were evaluated. Semistructured focus group meetings were conducted in 2022. Two focus groups were focused on sexually transmitted infection test packages and 2 were focused on prevention test packages. Data analysis was performed according to the principles of the Framework Method. The Consolidated Framework for Implementation Research was used to categorize the facilitators and barriers.

Results: In total, 19 participants, with a mean age of 34.32 (SD 14.70) years, participated in the focus groups. Important barriers were a lack of privacy information, too much and difficult information, and a commercial appearance. Important facilitators were the right amount of information, the right kind of tests, and the involvement of a health care professional. The need for a service such as Directlab Online was to ensure its availability for users' health and to maintain their health.

Conclusions: According to the participants, facilitators and barriers were comprehension of the information, the goal of the website, and the overall appearance of the service. Although the service was developed in cocreation with health care professionals and users, the needs did not align. The users preferred understandable and adequate, but not excessive, information. In addition, they preferred other types of tests to be available on the service. For future research, it would be beneficial to focus on cocreation between the involved medical professionals and users to develop, improve, and implement a service such as Directlab Online.

(JMIR Hum Factors 2024;11:e45115) doi:10.2196/45115

KEYWORDS

eHealth; usability; self-management; diagnostic test service; diagnostic; testing; test service; perspective; focus group; user need; user testing; implementation; qualitative; test result; laboratory test; laboratory result

XSL•FO
RenderX

Introduction

Background

Society is changing, and the world is becoming increasingly digital [1]. Health care lags in digital transformation, despite the potential of technology to improve the well-being of individuals [1,2]. The COVID-19 pandemic accelerated the development and use of technology in health care, also referred to as eHealth, with more digital consultations and increased use of home monitoring [3,4]. Furthermore, the COVID-19 pandemic, among others, has increased the need and willingness of individuals to perform self-management [5-7]. In patients with chronic diseases, self-management strategies are often used to support patients in dealing with treatment and lifestyle changes [8]. In addition, self-management strategies can be used to support individuals with home diagnostic tests [9]. The concept of self-management aligns with this positive health definition: "health as the ability to adapt and self-manage in the face of social, physical, and emotional challenges" [10,11].

eHealth can be used in 3 stages of laboratory diagnostic testing. The first stage is triage and advice on diagnostic testing, the second stage is the testing itself (ie, at home or a facility), and the third stage is the communication of the test results to the user. A systematic review by Versluis et al [9] showed that web-based diagnostic testing services were positively evaluated and preferred over clinic-based testing. However, most of the evaluated services only offered tests to detect sexually transmitted infections (STIs) [9].

eHealth services can support self-management, for example, with web-based services that support behavior and lifestyle changes (eg, Liva Healthcare) [12] and with websites where individuals can obtain health information (eg, Thuisarts.nl) [13]. In addition, there are multiple apps to support patients with chronic conditions such as hypertension, diabetes, or lower back pain [14-16].

In the Netherlands, a web-based service called Directlab Online (Saltro, part of Unilabs) offers individuals direct access to laboratory diagnostic tests independent of a health care provider [17]. It is a so-called direct-to-consumer platform. Directlab Online gives individuals direct digital access to diagnostic testing based on a triage that aligns with medical guidelines. Unlike the services identified in the systematic review by Versluis et al [9], Directlab Online offers a variety of diagnostic tests, for example, diagnostic tests for STIs, COVID-19, and vitamin deficiencies, as well as testing for health-related questions concerning fatigue and the prevention of heart disease. The results and the information on the website can give individuals insight into their health, which could support and motivate them to adopt healthier behaviors [12]. In addition, it supports users to be better informed about their health without the interference of a health care professional, which can lead to more efficient and accessible care [18]. Packages to test the health of individuals align with the patient-centered care approach, which can lead to a better quality of care [19]. Patient-centered care aims to empower patients to take charge of their health and actively participate in their health care [20]. Another term used is person-centered care, which is similar but does not solely focus on disease-related aspects, aligning better with the positive health definition [21].

To completely harness the potential and significance of Directlab Online, prioritizing high-quality and user-friendly service is paramount. Delving into the barriers and facilitators individuals encounter when using the service could provide invaluable insights, facilitating the enhancement of its user-friendliness and effectiveness. For example, known factors in dermatology that could influence the uptake of a digital service are, among others, financial aspects and accessibility for a digital service [22]. In a study by Vergouw et al [23], facilitators of and barriers to digital services for older adults in primary care were researched. Nonfamiliarities with web-based environments appeared to be a barrier, and efficiency was seen as an important facilitator for using a digital service in primary care [23]. In the review of STI testing by Versluis et al [9], concerns regarding complicated language and data handling insecurities were also discovered for ordering an STI test on the web. To our knowledge, no research has been conducted on facilitators, barriers, and needs of a direct-to-consumer platform that offers direct access to multiple diagnostic tests and web-based results. Identifying the needs, facilitators, and barriers will help determine what is necessary to optimize the use and improve the implementation of those services. This can give insight into the potential future directions for developing such services.

Objectives

This study aims to identify the facilitators of and barriers to using a service such as Directlab Online and to identify the needs regarding direct digital access to diagnostic testing. To achieve this, focus groups were conducted. Half of the focus groups focused on STI test packages and the other half on prevention test packages. STI tests and prevention test packages are the most ordered test packages on Directlab Online. The focus is on potential users, that is, those who have not used Directlab Online before, because we are interested in capturing people's first impression of the service.

Methods

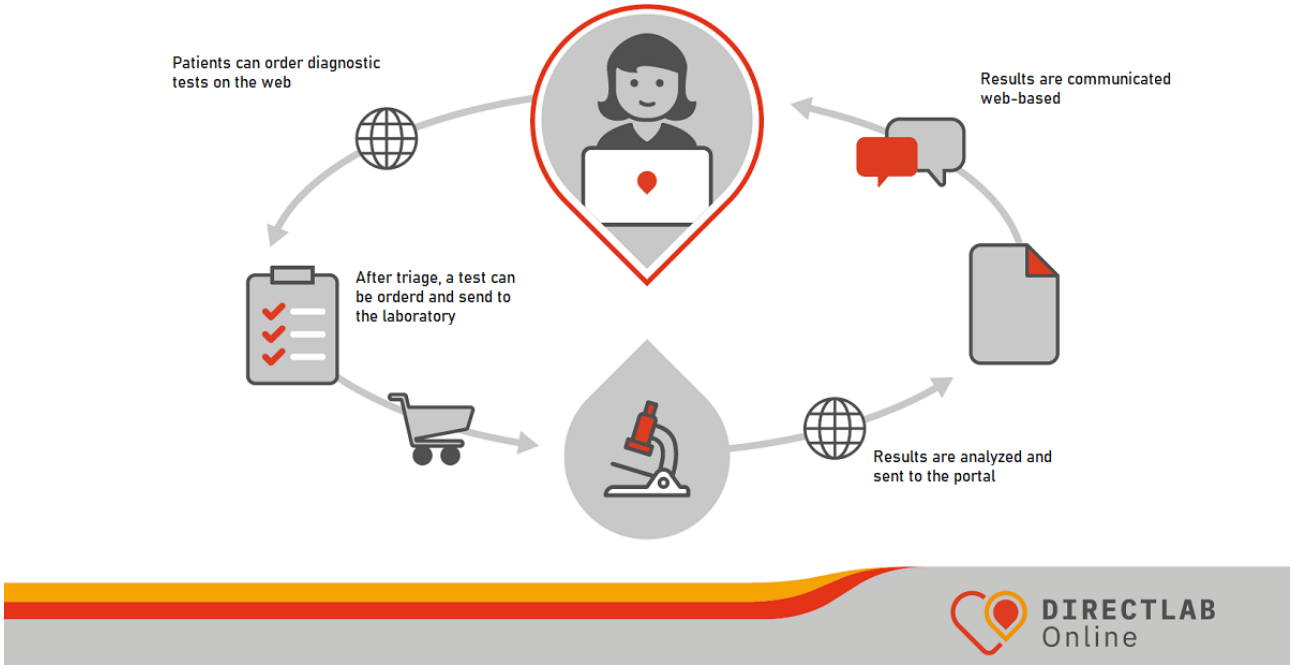
The Service: Directlab Online

Directlab Online is a Dutch, web-based service available for everyone, through which diagnostic tests can be ordered on the web [17]. The service was developed by a multidisciplinary innovation team of a diagnostic company (Saltro, part of Unilabs) and was launched in 2016 [24,25]. The process of using the service is presented in Figure 1. First, individuals undergo a web-based triage, based on medical guidelines, to determine whether the diagnostic tests are relevant and, if applicable, which tests are relevant. Second, individuals can order and buy associated tests. Depending on the diagnostic tests ordered, a self-sampling kit is sent to the individual's home address, or an appointment is scheduled at a blood collection center or a laboratory for collecting a blood sample. Once the laboratory receives the collected specimen, high-quality analyses are conducted. The results of the tests are communicated through a web-based, secure patient portal. Furthermore, deviating results are communicated to the patient's general practitioner but only if the patient has authorized it. The triage was based

XSL•FO
RenderX

on medical guidelines, and the diagnostic test packages were developed in cocreation with general practitioners and tested by them and laboratory specialists referred to as health care professionals. Diagnostic test packages consist of different parameters for diagnostic testing. For example, a test package for cholesterol measures the following parameters: low-density lipoproteins, high-density lipoproteins, triglycerides, and total cholesterol. Multimedia Appendix 1 provides a complete overview of the test packages that could be ordered on Directlab Online during the focus group meetings. Table 1 provides an overview of the prevention and STI test packages that were part of the discussions with the focus groups.

Figure 1. Stages of using Directlab Online.



XSL•FO
RenderX

Table 1. Test packages that are available on Directlab Online.

Category	Parameters
Prevention tests	
Health checkup	Check total cholesterol[a], low-density lipoproteins (LDL)[a], high-density lipoproteins (HDL)[a], triglycerides[a], HbA_{1c}[a], and albumin and creatinine ratio[b]
Health checkup at home[c]	Measure parameters via self-sampling of blood: total cholesterol[d], LDL[d], HDL[d], triglycerides[d], HbA_{1c}[d], and albumin/creatinine ratio[b]
Cholesterol	Check total cholesterol[a], LDL[a], HDL[a], and triglycerides[a]
Cholesterol at home[c]	Measure parameters via self-sampling of blood: total cholesterol[d], LDL[d], HDL[d], and triglycerides[d]
Anemia	Check hemoglobin[a], mean corpuscular volume[a], ferritin[a], and C-reactive protein[a]
Diabetes	Check glucose[a] and HbA_{1c}[a]
Healthy bones[c]	Check calcium[a] and vitamin D[a]
Healthy kidneys[c]	Check creatinine[a], glomerular filtration rate[a], and albumin/creatinine ratio[b]
Thyroid check	Check thyroid function via thyroid-stimulating hormone[a] and free T4[a]
Sexually transmitted infection tests	
Chlamydia	Check for chlamydia[e]
Gonorrhea	Check for gonorrhea[e]
HIV	Check for HIV[a]
Syphilis	Check for syphilis[a]
Hepatitis B	Check for hepatitis B[a]

[a]Blood sample needed for diagnostics; HbA_{1c}: hemoglobin A_{1c}.

[b]Urine sample needed for diagnostics.

[c]These tests are not available any more on Directlab Online after the service update.

[d]Blood sample needed for diagnostics collected by self-sampling.

[e]Oral, anal, vaginal, or urine sample needed for diagnostic tests.

Study Design and Participants

Focus group meetings were conducted with potential users of the service. As the Directlab Online service offers a wide variety of test packages, we focused on 2 specific categories (ie, prevention test and STI test packages). These test packages were ordered most frequently. Half of the focus groups focused on STI test packages, and the other half focused on the prevention test packages. The general inclusion criteria for the focus groups were speaking Dutch and not having used Directlab Online earlier. In addition, there were specific inclusion criteria to ensure that the sociodemographic characteristics of the participants in the focus groups were consistent with the characteristics of the target population of the test packages. Notably, a specific inclusion criterion for the focus group about STI testing was that the participants were aged between 18 and 30 years. The specific inclusion criterion for the focus groups about prevention test packages was that the participants were aged between 18 and 65 years. It is important to note that there were no specific health or disease requirements to participate. Focus group meetings were held until data saturation was reached.

Ethical Considerations

The study was declared to not fall within the scope of the Dutch Medical Research Involving Human Subjects Act by the Leiden University Medical Center Medical Ethics Committee (N21.101).

Procedure and Data Collection

The recruitment period started on October 25, 2021, and lasted until February 20, 2022. Participants were recruited via different web-based channels (eg, LinkedIn [LinkedIn Corporation] and Facebook [Meta platforms, Inc]). Individuals were invited to contact KS via email when interested. Then, KS sent them more information. In addition, questions were asked regarding their birth year and if they could understand Dutch. A few date options for web-based meetings were sent if the individual met the inclusion criteria. When individuals could participate, they received an email with the date and time, a link to the Zoom (Zoom Video Communications) platform where the meeting would be conducted (on the web), and a link to a web-based informed consent form, which they were asked to sign before participation. All participants had the right to withdraw at any moment. The focus group meetings occurred between January

XSL•FO
RenderX

10 and March 2, 2022, in the presence of MH and KS [26]. KS led the focus groups, and MH managed the time and assisted with technical issues. The focus group meetings were in a semistructured format, following a predefined topic list with open-ended questions to leave space for discussion (Multimedia Appendix 2). First, general questions were asked regarding using eHealth to see how familiar participants were with eHealth. Second, participants were provided 10 minutes to view the website of Directlab Online and navigate through the website on a computer or mobile phone; no further instructions were given. When time was up, questions were asked regarding the website in general (eg, the first impression, whether they needed help when using the website, and whether they found the website attractive). While navigating the website, they had the option to write down notes or vocalize their impressions, expressing their observations, preferences, and feelings about the website [27]. Third, participants were instructed to go through Directlab Online, do some triages, and look at their test advice. Notably, we allowed participants to navigate through the process as normal users would. Hence, they were required to peruse informational materials, undergo a triage process involving medical inquiries concerning their symptoms, and obtain guidance regarding testing. Subsequently, questions were asked regarding the triage service, facilitators of and barriers to using Directlab Online, and the participants' needs for such a service. At the end of the focus groups, they received a digital gift card of €25 (US $27).

Data Analysis

All focus group meetings were audio recorded for subsequent analyses and were transcribed (intelligent) verbatim. When the transcripts were completed, the audio records were deleted. Two reviewers, MH and KS, conducted the qualitative data analysis according to the principles of the Framework Method [28]. The Framework Method is a systematic and flexible approach commonly used for the thematic analysis of health research semistructured interview data [29]. The method combines deductive and inductive techniques, which align with the aim of the study to identify specific issues regarding the use of Directlab Online and leave space to identify needs and opportunities that have not been formulated a priori. First, open coding was performed independently by the 2 reviewers, KS and MH. The interview data were coded using the software Atlas.ti 22 (Atlas.ti 22 Scientific Software Development). Second, the codes were compared between the 2 reviewers. Third, the codes were grouped into categories, resulting in the analytical framework. Fourth, for identifying the facilitators and barriers, the Consolidated Framework for Implementation Research (CFIR) was used [30]. The framework is widely used for the content analysis of qualitative data regarding the factors influencing implementation success [30]. Furthermore, the framework is comprehensive and makes it convenient to systematically study a wide array of facilitators and barriers [31]. In addition, using this framework made it possible to compare the findings and transfer them to other implementation studies [32]. The CFIR is a theory-driven model and comprises five domains: (1) the innovation domain, (2) the outer setting domain, (3) the inner setting domain, (4) the individuals' domain, and (5) the implementation process [30,33]. Identified facilitators and barriers were placed within the CFIR domains. Final themes were achieved via discussion and consensus between researchers KS and MH.

Results

Participant Characteristics

Data saturation was reached after forming 4 focus groups with 19 participants. The characteristics of the participants are shown in Table 2. The participants were aged 20 to 61 (mean 34.32, SD 14.70) years. The number of male (9/19, 47%) and female (10/19, 53%) participants was almost equal. The focus group meetings lasted around 90 minutes per group.

Age differed over the 2 different focus groups, as aligned with the target population of the diagnostic test packages. Overall, the experiences and choices of the focus groups regarding the website were the same. Therefore, in most cases, the focus group results were discussed together. When the results differed between the 2 groups, this was specified. Different themes around usability, facilitators, barriers, and needs emerged from the data and are elaborated in subsequent sections.

XSL•FO
RenderX

Table 2. Characteristics of the participants (N=19).

Participant	Gender	Age (years)	Focus group[a]
1	Woman	27	1
2	Woman	25	1
3	Man	24	1
4	Man	30	1
5	Woman	20	1
6	Woman	25	2
7	Woman	46	2
8	Woman	59	2
9	Man	24	2
10	Man	20	2
11	Woman	25	3
12	Man	25	3
13	Woman	30	3
14	Man	24	3
15	Man	39	4
16	Woman	58	4
17	Woman	59	4
18	Man	30	4
19	Man	62	4

[a]Groups 1 and 3 focused on sexually transmitted infection packages, and groups 2 and 4 focused on prevention packages.

Facilitators of and Barriers to the Uptake of Innovation

The identified barriers and facilitators were categorized specifically into the following 3 CFIR domains: innovation domain, outer setting domain, and individuals domain. The other 2 domains of the CFIR (ie, inner setting and implementation process) did not align with the facilitators and barriers mentioned by the participants and were therefore not discussed. Table 3 provides insight into the most essential and changeable facilitators and barriers identified. Therefore, it is not an exhaustive list of all potential barriers and facilitators that influenced the service uptake. It is notable that certain factors can be considered as a facilitator and barrier. For example, financial costs are frequently mentioned as a factor affecting the willingness to use digital health services [33]. When there are high user costs, it is a barrier; however, low costs can be considered a facilitator. The identified facilitators and barriers are explained in detail and explained per domain in subsequent sections.

XSL•FO
RenderX

Table 3. Facilitators and barriers derived from the focus groups embedded in the Conceptual Framework for Implementation Research (CFIR).

Domain of CFIR	Domain description	Results
Innovation domain		
Innovation source	The group that developed and visibly sponsored the use of the innovation is reputable, credible, and trustable.	• The general practitioner group that developed and visibly sponsored the service was reputable, credible, and trustable, which resulted in a reliable service. • Information about privacy and presenting good reviews improved reliability and credibility. • Commercial appearance influenced the credibility. Furthermore, stock pictures influenced the credibility.
Innovation relative advantage	The innovation is better than other available innovations or current practices.	• The service was easy to use, which made the service accessible. • It was easy to use the service without visiting the general practitioner.
Innovation complexity	The innovation is complicated, which may be reflected by its scope and the nature and number of connections and steps.	• Too many testing possibilities and too much information made the website less user-friendly. • The search bar and filters on the website increased the user-friendliness of the website. • Using multiple medical words made the service difficult to comprehend.
Outer setting domain		
Partnerships and connections	The inner setting is networked with external entities, including referral networks, academic affiliations, and professional organization networks.	• The service was linked with academic institutions and other medical professionals, which increased the reliability of the service for users.
Societal pressure	Mass media campaigns, advocacy groups, or social movements or protests drive the implementation and delivery of the innovation.	• Media campaigns, reviews, and blogs could help stimulate participants to use the service.
Individuals domain: subdomain patient characteristics		
Capability	The individual has interpersonal competence, knowledge, and skills to fulfill a "Role" (different characteristics of individuals).	• If participants had experience with a similar service, they felt more confident in using the service. Otherwise, feelings of anxiety or tension could have influenced their competence, knowledge, and skills.

Facilitators and Barriers in the Innovation Domain

Innovation Source

Participants mentioned different factors that were related to the innovation source of the innovation domain. These factors mainly influenced the credibility and trustworthiness in a positive (ie, facilitator) or negative (ie, barrier) way. First, the website's commercial appearance were the most frequently mentioned barriers that influenced its reliability. For example, participants mentioned that the option to buy a gift card for a diagnostic test package did not align with a website designed for health. In addition, regarding the high prices for diagnostic test packages and the website's general appearance, they said the following:

> *The website said: buy this. But I want to know why this test?* [Participant 4]
>
> *I found it a very commercial website; this lowers my enthusiasm.* [Participant 8]

Participants did not notice that health care professionals were involved in the service and partly developed the service, while this could increase the credibility of the website.

Second, the availability of reviews was frequently mentioned as a facilitator for reliability and credibility but as a barrier in some cases. Good reviews could be considered as a facilitator, and bad reviews could be considered as a barrier to experiencing the website as reliable and trustworthy. The following was said about this view:

XSL•FO
RenderX

> *Yes, [...] I found it important if I go to a new website to sell or buy something to see that others used the site and what they bought.* [Participant 13]

Third, 37% (7/19) of the participants mentioned the facilitator's "privacy." For the participants, it was important to know where the data were stored and for how long. However, this information was difficult to find on the website:

> *And then it is the question of how long data is stored and how that is important to know.* [Participant 8]
>
> *I want to know, what happens to the data and how long is it stored?* [Participant 16]

Participant 7 pointed out that a clear and transparent privacy statement could be a unique selling point of the service.

Finally, the most mentioned barrier in the innovation source was the presence of stock pictures on Directlab Online:

> *[...] those stock pictures on the website; they gave an image of unreliability.* [Participant 3]

Participants mentioned that pictures of real people or famous people who used the tests could be a facilitator and positively influence the reliability and use of the service. Furthermore, they mentioned that a short video with education and instructions about diagnostic test packages could improve the triage's clarity and the diagnostic packages' content.

Innovation Relative Advantage

Participants mentioned several factors regarding why they would use this innovative service. These factors were mostly related to the accessibility of the service compared to other services or normal practice. For example, the easiness of ordering a test on the web without going to the general practitioner was a relative advantage of the service, as mentioned by a participant:

> *Yes, I would rather order online because going to the general practitioner [...] it takes time.* [Participant 7]

Furthermore, another participant mentioned the benefit of ordering a test on the web without going to the general practitioner:

> *Hmm yes, I thought of a few things when I first saw the website... of the vitamin tests, STI tests, and COVID tests, I thought yes, you do not want to go to the general practitioner for that. Especially for STI testing, the threshold is high. In this way, you still test and see if you are healthy.* [Participant 1]

However, the relative advantage was negatively influenced by the high costs of the tests. One participant stated:

> *The costs will stop people from buying anything.* [Participant 17]

Innovation Complexity

Several facilitators and barriers that influenced the complexity of the service were mentioned by the participants.

First, the number of test packages and parameters available was confusing. It became clear from the focus groups that offering the "right" number of diagnostic tests was important; participants were not enthusiastic about a test package with many separate parameters. Participants mentioned that they were optimistic about the possibility of ordering STI testing, COVID-19 testing, and some prevention tests. However, they mentioned that after the triage, they received advice to select many different test packages. Recommending many diagnostic test packages to the participants was a barrier because they were confused about which test package was important for them. Furthermore, the high amount of information provided about these test packages was experienced as challenging by approximately half (11/19, 58%) of the participants:

> *When I open the website, a lot of information is present. Too many tests are available. Of course, this website wants to sell tests, but... I do not know. I found the home page too complicated, too unclear.* [Participant 13]

Second, the language used on the website was a factor that influenced the use of the service. The language on the home page was experienced as straightforward and was therefore a facilitator. However, when completing the triage and choosing the diagnostic package, the information was more challenging to understand. Notably, medical and incomprehensible terms were used:

> *I think you have a very broad target group of people who would like to use this, and I think it is written for the somewhat well-educated, reasonably well-informed citizen, shall we say [...] Offer more comfort to people by using less difficult vocabulary.* [Participant 8]

Third, the participants mentioned elements of the website that influenced user-friendliness. Participants were happy with the filters in the search bar to find a particular test, as well as the search function and the website's colors:

> *Personally, I found the website easy to use, and what I experienced as very positive were the filters [...].* [Participant 14]

However, approximately one-third (6/19, 32%) of the participants found the website unclear—among others, due to too much text—and complicated (eg, where to find what they were looking for), and they found the home page too busy.

Facilitators and Barriers in the Outer Setting Domain

Partnerships and Connections

The service was linked to academic institutions, which increased its reliability. Mentioning partners would increase the uptake according to the participants:

> *Yes, mentioning partners would be nice. And famous names always attract attention.* [Participant 13]

Societal Pressure

Participants mentioned that reviews and blogs could help in increasing the use of the service and its reliability:

> *You want to read reviews and experiences of others.* [Participant 5]

XSL•FO
RenderX

Facilitators and Barriers in the Individuals Domain

The individual's skills and knowledge regarding services such as Directlab Online influenced their willingness to use the service and their perception of potentially using it. The younger participants (aged 20 to 30 years) mentioned that they had experience with this type of website, which reassured them to use this service. However, some older participants (aged ≥39 years) had less experience with digital services in general and mentioned some anxiety and tension when they needed to order a test. Some (4/19, 21%) preferred to go to the general practitioner for diagnostic tests. However, participants of all age groups mentioned the benefit of ordering STI tests on the web without visiting the general practitioner.

Future Needs

Different needs were identified regarding services such as Directlab Online. First, the service's purpose must be more explicit for the participants. For them, it was unclear whether the service could help them self-manage their health:

> *And this is what I miss on the website; what is in it for me and my health as a patient or consumer?* [Participant 19]

Second, there was a need to understand the advantages of ordering diagnostic tests on the web (eg, more accessible than going to the general practitioner for tests). Participants wanted this information to be more evident on the website. Third, after receiving their results, the participants explained that they preferred to have more information regarding how they could remain healthy or what they could do to become healthier. It could help, according to the participants, to let them know more specifically that general practitioners make the diagnostic test packages designed for the service. All participants saw the benefit of ordering STI diagnostic test packages on the web and undergoing them at home. The current offer of diagnostic test packages does not meet the wishes of all participants. There was a need for additional tests, such as tests for food allergies, testosterone levels, fertility, or urinary infections. A participant mentioned as follows:

> *I want a urine tract infection test; those are relatively cheap, I think [...].* [Participant 1]

Discussion

Principal Findings

This qualitative study aimed to evaluate the facilitators of and barriers to web-based direct access to diagnostic test services from the perspective of potential users. In addition, the study tried to identify the need to use such services. The study showed that a tailored amount of information could benefit the service. Participants need to use a service such as Directlab Online to ensure that the website is available for their health. The participants needed to see the benefit of a diagnostic test package. Identified barriers and facilitators were categorized using the CFIR. The study showed that a lack of privacy, information overload, and a commercial appearance were important barriers. Facilitators included providing the right amount of information on the service and involving a health care professional in developing the service. In addition, the study showed that a tailored amount of information could benefit the use of the service. In short, we noticed that several facilitators and barriers were influencing the reliability or accessibility of the service. For example, the commercial appearance and lack of privacy information contributed to a less reliable service for the potential users, and ordering a test on the web without a health care professional influenced the accessibility.

Directlab Online is a service for users to support themselves in self-managing their health. An important quality-enhancing element for Directlab Online was that health care professionals were actively involved in developing this service. Health care professionals significantly influenced the content and information shown on the website. However, the focus groups with potential users identified needs and wishes that did not completely align with the ideas of the general practitioner. To illustrate, health care professionals preferred other types of diagnostic test packages on the web than those that the participants preferred to use. Furthermore, the general practitioners preferred detailed information on the website, whereas this information overload was not always beneficial for the participants. A study by Talboom-Kamp et al [34] regarding a web-based results portal discussed the complex balance between the general practitioner's necessities and participants' needs for the right amount of understandable information. Presenting information requires a balance between an overload of medical information and the information users need to understand test packages and results. A potential way to solve overwhelming participants with information is to not present all the information directly in one view to the participant but by offering clickable links or short videos [34].

This study used the CFIR to identify and categorize the facilitators and barriers. In another study, Versluis et al [33] performed an inventory to determine the obstacles that must be overcome and how to optimize eHealth in primary care using this framework. They found similar results to our study; costs and privacy issues were identified as important barriers. In addition, in line with other studies, the following facilitators were identified as having "experience with eHealth" and "easiness of use" [33,35]. In comparison with other studies using the CFIR to classify facilitators and barriers, similar factors were predominantly identified. A notable factor highlighted in the study by Verweij et al [36] involving patients with cancer using a digital self-monitoring system was the necessity to elucidate the service's added value, alongside concerns regarding privacy issues. However, other factors were also mentioned, such as the connection with health care professionals, which were not identified in our study. The target population (patients with cancer) could be an important explanation for this difference. The comparison with other literature revealed that, irrespective of the type of digital service or the user population, the facilitators and barriers remained quite consistent. This study's inventory could help determine what obstacles need to be overcome and how we might optimize an application such as Directlab Online.

Depending on the participants, mainly influenced by age, some would use a web-based website to organize their health. In contrast, other participants, mainly older participants, were more

XSL•FO
RenderX

at ease with visiting the general practitioner and organizing their health directly via the general practitioner [37]. In this study, the older participants preferred to visit the general practitioner, which could lead to the cautious conclusion that web-based direct access to diagnostic services is not attractive to everyone [37]. In addition, this study showed that using a service such as Directlab Online is not only related to age but also related to the user's health problem and that the type of test package was important. Participants' needs were to feel the relevance of ordering a diagnostic test package on the web instead of visiting the general practitioner. The relevance was clear for the STI test packages but unclear for other diagnostic test packages. The study results showed that it remains important to involve all end users in the service to ensure that the service supports the needs of the target population [38]. Directlab Online was developed with general practitioners, and the elements that they found important were integrated into the service. While this study provided insight into the facilitators and barriers of potential users, it appeared that these things were not the same. It is important for a reliable and proper service that both perspectives of all stakeholders should be included in (further) development of such services. Finally, the facilitators and barriers to using a service such as Directlab Online found in this study could be used to optimize this service and other comparable services.

Strengths and Limitations

There is a lot of direct access to diagnostic testing services available, mainly when it entails STI diagnostic test packages. However, not many of them have a scientific basis or are developed by medical professionals. This is the first study that examined the facilitators of and barriers to a service that provides more diagnostic test packages than only STI tests and which is developed in cocreation with health care professionals. Another strength of the study was that the CFIR was used to analyze the facilitators and barriers mentioned in the focus groups. Embedding the facilitators and barriers in this framework made the comparison with other research easier. In addition, the domains identified by the CFIR can help to find the right implementation strategy [33,39].

This study focused on potential users because we were interested in their first impression of the service. The rationale was that, in the real world, such a service could be visited by many new users [40]. Previous experiences have not biased the impression of potential users. However, this could also be a limitation because participants who did use Directlab Online before could have another opinion regarding the service. This made the results less generalizable. Another limitation is that the mean age of participants was relatively low, making it more difficult to generalize the results to the general Dutch population. However, all participants, independent of age, mentioned the benefit of ordering STI tests on the web. The service showed benefits for participants who were ashamed to visit a general practitioner for a diagnostic package and for participants who wished to order tests in an accessible, nonbinding manner.

Future Research

Directlab Online is a service developed for a wide range of users. However, this study showed that it is important to include end users to ensure that the service aligns with the population's needs. Cocreation with end users and medical professionals could be a solution to solve disbalances in wishes and needs between them and to improve an eHealth application [38]. For future research, organizing cocreation sessions and analyzing their results could be beneficial to improve the service. Finally, in future research, information about the influence of the diagnostic test's result on the user's lifestyle could be analyzed. Namely, this could result in a preventive role for a service such as Directlab Online to improve the health of a population.

Conclusions

According to participants, information provision, comprehension, and the overall appearance of the website were the most important elements that influenced the use and uptake of a direct-to-consumer website for diagnostic test packages. Barriers, such as the commercial appearance and lack of privacy information, negatively influenced reliability and accessibility. The study showed that it is important to include relevant stakeholders in creating an eHealth intervention because there was a disbalance between the users' needs and what the involved general practitioners considered necessary. Future research could take a quantitative approach to further identify the needs regarding test packages and to identify the demographics of users and the influence of test results on the behavior of users. Directlab Online offers opportunities for more web-based self-management of health.

Acknowledgments

The authors would like to thank all participants for their input and contributions during the focus groups. Funding for this research was provided by Unilabs.

Conflicts of Interest

During the research, KS and EPWATK were employees at Unilabs.

Multimedia Appendix 1
Diagnostic test packages overview of Directlab Online.
[DOCX File , 20 KB - humanfactors_v11i1e45115_app1.docx]

Multimedia Appendix 2

XSL•FO
RenderX

Semistructured interview protocol.
[DOCX File , 15 KB - humanfactors_v11i1e45115_app2.docx]

References

1. Chui M, Roberts R, Yee L. McKinsey technology trends outlook 2022. McKinsey & Company. 2022 Aug. URL: https://tinyurl.com/3tv93phy [accessed 2024-04-28]
2. Rawstron E, Baulderstone M. Plugging healthcare workers into the digital future. KPMG. 2022. URL: https://tinyurl.com/bdm4dpu9 [accessed 2024-04-28]
3. Clipper B. The influence of the COVID-19 pandemic on technology: adoption in health care. Nurse Lead 2020 Oct;18(5):500-503 [FREE Full text] [doi: 10.1016/j.mnl.2020.06.008] [Medline: 32837346]
4. Keuper J, Batenburg R, Verheij R, van Tuyl L. Use of e-health in Dutch general practice during the COVID-19 pandemic. Int J Environ Res Public Health 2021 Nov 26;18(23):12479 [FREE Full text] [doi: 10.3390/ijerph182312479] [Medline: 34886204]
5. Menting J, Schelven F, Boeije H. Gevolgen van de coronapandemie voor gezondheid, behandeling en zelfmanagement van mensen met een chronische ziekte. Nivel. 2020. URL: https://nivel.nl/nl/publicatie/gevolgen-van-de-coronapandemie-voor-gezondheid-behandeling-en-zelfmanagement-van-mensen [accessed 2024-04-28]
6. Flinterman L, Groenewegen P, Verheij R. Zorglandschap en zorggebruik in een veranderende eerste lijn. Nivel. 2018. URL: https://www.nivel.nl/nl/publicatie/zorglandschap-en-zorggebruik-een-veranderende-eerste-lijn [accessed 2024-04-28]
7. Gauthier-Beaupré A, Battistini BJ, Kuziemsky C, Jutai JW. Policymaker perspectives on self-management of disease and disabilities using information and communication technologies. Health Res Policy Syst 2023 Jun 14;21(1):52 [FREE Full text] [doi: 10.1186/s12961-023-01004-7] [Medline: 37316838]
8. McGowan P. Self-management: a background paper. University of Victoria - Centre on Aging. 2005. URL: https://tinyurl.com/378my7xc [accessed 2024-04-28]
9. Versluis A, Schnoor K, Chavannes NH, Talboom-Kamp EP. Direct access for patients to diagnostic testing and results using eHealth: systematic review on eHealth and diagnostics. J Med Internet Res 2022 Jan 12;24(1):e29303 [FREE Full text] [doi: 10.2196/29303] [Medline: 35019848]
10. Huber M, Knottnerus JA, Green L, van der Horst H, Jadad AR, Kromhout D, et al. How should we define health? BMJ 2011 Jul 26;343(jul26 2):d4163. [doi: 10.1136/bmj.d4163] [Medline: 21791490]
11. Huber M, van Vliet M, Boers I. Heroverweeg uw opvatting van het begrip 'gezondheid'. Ned Tijdschr Geneeskd 2016;160:A7720 [FREE Full text]
12. Komkova A, Brandt CJ, Hansen Pedersen D, Emneus M, Sortsø C. Electronic health lifestyle coaching among diabetes patients in a real-life municipality setting: observational study. JMIR Diabetes 2019 Mar 12;4(1):e12140 [FREE Full text] [doi: 10.2196/12140] [Medline: 30860486]
13. Spoelman WA, Bonten TN, de Waal MW, Drenthen T, Smeele IJ, Nielen MM, et al. Effect of an evidence-based website on healthcare usage: an interrupted time-series study. BMJ Open 2016 Nov 09;6(11):e013166 [FREE Full text] [doi: 10.1136/bmjopen-2016-013166] [Medline: 28186945]
14. Machado GC, Pinheiro MB, Lee H, Ahmed OH, Hendrick P, Williams C, et al. Smartphone apps for the self-management of low back pain: a systematic review. Best Pract Res Clin Rheumatol 2016 Dec;30(6):1098-1109. [doi: 10.1016/j.berh.2017.04.002] [Medline: 29103552]
15. El-Gayar O, Timsina P, Nawar N, Eid W. Mobile applications for diabetes self-management: status and potential. J Diabetes Sci Technol 2013 Jan 01;7(1):247-262 [FREE Full text] [doi: 10.1177/193229681300700130] [Medline: 23439183]
16. Alessa T, Hawley MS, Hock ES, de Witte L. Smartphone apps to support self-management of hypertension: review and content analysis. JMIR Mhealth Uhealth 2019 May 28;7(5):e13645 [FREE Full text] [doi: 10.2196/13645] [Medline: 31140434]
17. Unilabs Online homepage. Unilabs Online. URL: https://directlabonline.nl/ [accessed 2024-04-24]
18. Holman H, Lorig K. Patient self-management: a key to effectiveness and efficiency in care of chronic disease. Public Health Rep 2004 Aug 02;119(3):239-243 [FREE Full text] [doi: 10.1016/j.phr.2004.04.002] [Medline: 15158102]
19. Bergeson SC, Dean JD. A systems approach to patient-centered care. JAMA 2006 Dec 20;296(23):2848-2851. [doi: 10.1001/jama.296.23.2848] [Medline: 17179462]
20. Reynolds A. Patient-centered care. Radiol Technol 2009;81(2):133-147. [Medline: 19901351]
21. Starfield B. Is patient-centered care the same as person-focused care? Perm J 2011;15(2):63-69 [FREE Full text] [doi: 10.7812/TPP/10-148] [Medline: 21841928]
22. Ariens LF, Schussler-Raymakers FM, Frima C, Flinterman A, Hamminga E, Arents BW, et al. Barriers and facilitators to eHealth use in daily practice: perspectives of patients and professionals in dermatology. J Med Internet Res 2017 Sep 05;19(9):e300 [FREE Full text] [doi: 10.2196/jmir.7512] [Medline: 28874336]
23. Vergouw JW, Smits-Pelser H, Kars MC, van Houwelingen T, van Os-Medendorp H, Kort H, et al. Needs, barriers and facilitators of older adults towards eHealth in general practice: a qualitative study. Prim Health Care Res Dev 2020 Dec 02;21:e54 [FREE Full text] [doi: 10.1017/S1463423620000547] [Medline: 33263272]
24. Over Saltro. Saltro. URL: https://saltro.nl/ [accessed 2024-04-28]

XSL•FO
RenderX

25. Unilabs homepage. Unilbas. URL: https://unilabs.com/ [accessed 2022-11-28]
26. Willemsen RF, Aardoom JJ, Chavannes NH, Versluis A. Online synchronous focus group interviews: practical considerations. Qual Res 2023;23(6):1810-1820. [doi: 10.1177/14687941221110161]
27. Bonten TN, Rauwerdink A, Wyatt JC, Kasteleyn MJ, Witkamp L, Riper H, et al. Online guide for electronic health evaluation approaches: systematic scoping review and concept mapping study. J Med Internet Res 2020 Aug 12;22(8):e17774 [FREE Full text] [doi: 10.2196/17774] [Medline: 32784173]
28. Spencer L, Ritchie J, Lewis J, Dillon L. Quality in qualitative evaluation: a framework for assessing research evidence. Government Chief Social Researcher's Office, United Kingdom. 2003 Aug. URL: https://assets.publishing.service.gov.uk/media/5a8179c1ed915d74e33fe69e/Quality-in-qualitative-evaulation_tcm6-38739.pdf [accessed 2024-04-28]
29. Gale NK, Heath G, Cameron E, Rashid S, Redwood S. Using the framework method for the analysis of qualitative data in multi-disciplinary health research. BMC Med Res Methodol 2013 Sep 18;13:117 [FREE Full text] [doi: 10.1186/1471-2288-13-117] [Medline: 24047204]
30. Breimaier HE, Heckemann B, Halfens RJ, Lohrmann C. The Consolidated Framework for Implementation Research (CFIR): a useful theoretical framework for guiding and evaluating a guideline implementation process in a hospital-based nursing practice. BMC Nurs 2015 Aug 12;14:43 [FREE Full text] [doi: 10.1186/s12912-015-0088-4] [Medline: 26269693]
31. Safaeinili N, Brown-Johnson C, Shaw JG, Mahoney M, Winget M. CFIR simplified: pragmatic application of and adaptations to the Consolidated Framework for Implementation Research (CFIR) for evaluation of a patient-centered care transformation within a learning health system. Learn Health Syst 2020 Sep 26;4(1):e10201 [FREE Full text] [doi: 10.1002/lrh2.10201] [Medline: 31989028]
32. Ross J, Stevenson F, Lau R, Murray E. Factors that influence the implementation of e-health: a systematic review of systematic reviews (an update). Implement Sci 2016 Oct 26;11(1):146 [FREE Full text] [doi: 10.1186/s13012-016-0510-7] [Medline: 27782832]
33. Versluis A, van Luenen S, Meijer E, Honkoop PJ, Pinnock H, Mohr DC, et al. SERIES: eHealth in primary care. Part 4: addressing the challenges of implementation. Eur J Gen Pract 2020 Dec;26(1):140-145 [FREE Full text] [doi: 10.1080/13814788.2020.1826431] [Medline: 33025820]
34. Talboom-Kamp E, Tossaint-Schoenmakers R, Goedhart A, Versluis A, Kasteleyn M. Patients' attitudes toward an online patient portal for communicating laboratory test results: real-world study using the eHealth impact questionnaire. JMIR Form Res 2020 Mar 04;4(3):e17060 [FREE Full text] [doi: 10.2196/17060] [Medline: 32024632]
35. Cajita MI, Hodgson NA, Lam KW, Yoo S, Han HR. Facilitators of and barriers to mHealth adoption in older adults with heart failure. Comput Inform Nurs 2018 Aug;36(8):376-382 [FREE Full text] [doi: 10.1097/CIN.0000000000000442] [Medline: 29742549]
36. Verweij L, Smit Y, Blijlevens NM, Hermens RP. A comprehensive eHealth implementation guide constructed on a qualitative case study on barriers and facilitators of the digital care platform CMyLife. BMC Health Serv Res 2022 Jun 06;22(1):751 [FREE Full text] [doi: 10.1186/s12913-022-08020-3] [Medline: 35668491]
37. Wilson J, Heinsch M, Betts D, Booth D, Kay-Lambkin F. Barriers and facilitators to the use of e-health by older adults: a scoping review. BMC Public Health 2021 Aug 17;21(1):1556 [FREE Full text] [doi: 10.1186/s12889-021-11623-w] [Medline: 34399716]
38. Kasteleyn MJ, Versluis A, van Peet P, Kirk UB, van Dalfsen J, Meijer E, et al. SERIES: eHealth in primary care. Part 5: a critical appraisal of five widely used eHealth applications for primary care - opportunities and challenges. Eur J Gen Pract 2021 Dec;27(1):248-256 [FREE Full text] [doi: 10.1080/13814788.2021.1962845] [Medline: 34432601]
39. Powell BJ, Waltz TJ, Chinman MJ, Damschroder LJ, Smith JL, Matthieu MM, et al. A refined compilation of implementation strategies: results from the expert recommendations for implementing change (ERIC) project. Implement Sci 2015 Feb 12;10:21 [FREE Full text] [doi: 10.1186/s13012-015-0209-1] [Medline: 25889199]
40. Grindrod KA, Li M, Gates A. Evaluating user perceptions of mobile medication management applications with older adults: a usability study. JMIR Mhealth Uhealth 2014 Mar 14;2(1):e11 [FREE Full text] [doi: 10.2196/mhealth.3048] [Medline: 25099993]

Abbreviations

CFIR: Consolidated Framework for Implementation Research
STI: sexually transmitted infection

XSL•FO
RenderX

Edited by A Kushniruk; submitted 30.01.23; peer-reviewed by S Choemprayong, Y Chu, CE Moe; comments to author 19.12.23; revised version received 02.01.24; accepted 20.03.24; published 10.05.24.

Please cite as:
Schnoor K, Talboom-Kamp EPWA, Hajtić M, Chavannes NH, Versluis A
Facilitators of and Barriers to the Use of a Digital Self-Management Service for Diagnostic Testing: Focus Group Study With Potential Users
JMIR Hum Factors 2024;11:e45115
URL: https://humanfactors.jmir.org/2024/1/e45115
doi:10.2196/45115
PMID:38728071

©Kyma Schnoor, Esther P W A Talboom-Kamp, Muamer Hajtić, Niels H Chavannes, Anke Versluis. Originally published in JMIR Human Factors (https://humanfactors.jmir.org), 10.05.2024. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on https://humanfactors.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Evaluating the Energy Efficiency of Popular US Smartphone Health Care Apps: Comparative Analysis Study Toward Sustainable Health and Nutrition Apps Practices

Abdullah Almasri[1*], PhD; Tatyana Y El-Kour[2*], RDN, PhD; Liliana Silva[3*], PhD; Yousef Abdulfattah[4*], BSc

[1]College of Humanities and Sciences, Prince Sultan University, Riyadh, Saudi Arabia
[2]Media Psychology Department, Fielding Graduate University, Santa Barbara, CA, United States
[3]Faculty of Health Sciences, University Fernando Pessoa, Porto, Portugal
[4]College of Computer & Information Sciences, Prince Sultan University, Riyadh, Saudi Arabia
[*]all authors contributed equally

Corresponding Author:
Abdullah Almasri, PhD
College of Humanities and Sciences
Prince Sultan University
PPP2+CGP, Nasser Bin Farhan Street
King Salman Neighborhood
Riyadh, 12435
Saudi Arabia
Phone: 966 4948000
Email: aalmasri@psu.edu.sa

Abstract

Background: The emergence of smartphones has sparked a transformation across multiple fields, with health care being one of the most notable due to the advent of mobile health (mHealth) apps. As mHealth apps have gained popularity, there is a need to understand their energy consumption patterns as an integral part of the evolving landscape of health care technologies.

Objective: This study aims to identify the key contributors to elevated energy consumption in mHealth apps and suggest methods for their optimization, addressing a significant void in our comprehension of the energy dynamics at play within mHealth apps.

Methods: Through quantitative comparative analysis of 10 prominent mHealth apps available on Android platforms within the United States, this study examined factors contributing to high energy consumption. The analysis included descriptive statistics, comparative analysis using ANOVA, and regression analysis to examine how certain factors impact energy use and consumption.

Results: Observed energy use variances in mHealth apps stemmed from user interactions, features, and underlying technology. Descriptive analysis revealed variability in app energy consumption (150-310 milliwatt-hours), highlighting the influence of user interaction and app complexity. ANOVA verified these findings, indicating the critical role of engagement and functionality. Regression modeling (energy consumption = $\beta + \beta_1 \times$ notification frequency + $\beta_2 \times$ GPS use + $\beta_3 \times$ app complexity + ε), with statistically significant *P* values (notification frequency with a *P* value of .01, GPS use with a *P* value of .05, and app complexity with a *P* value of .03), further quantified these bases' effects on energy use.

Conclusions: The observed differences in the energy consumption of dietary apps reaffirm the need for a multidisciplinary approach to bring together app developers, end users, and health care experts to foster improved energy conservation practice while achieving a balance between sustainable practice and user experience. More research is needed to better understand how to scale-up consumer engagement to achieve sustainable development goal 12 on responsible consumption and production.

(JMIR Hum Factors 2024;11:e58311) doi:10.2196/58311

KEYWORDS

mobile health; energy consumption in health care smartphone apps; dietary tracking apps; optimization and sustainability in mobile health; user engagement and experience; Android apps performance; digital health technologies; app; apps; applications; digital health; energy; consumption; sustainable; sustainability; environment; environmental; use; smartphone; smartphones; electricity; electrical; mobile phone

XSL•FO
RenderX

Introduction

Background

Nations worldwide and researchers from various disciplines are increasingly focusing on sustainable and energy-efficient techniques for energy production. The works of Bhaskar et al [1], Muthanna et al [2], and Ashfaq et al [3] exemplified the innovative approaches being developed in this domain, highlighting the significance of renewable energy applications, unmanned aerial vehicle path scheduling in the Internet of Things, and secure energy trading with machine learning and blockchain technology, respectively. In today's health care scene, smartphones stand as crucial companions, seamlessly connecting the realms of technology and wellness promotion. The surge in popularity of mobile health (mHealth) apps reflects a broader movement toward adopting energy-smart habits in all facets of mobile computing. This trend underscores the pivotal role of crafting sustainable software to lessen our ecological footprint, a goal echoed by the strides made in green computing and energy-saving innovations [4-6]. These apps mark a transformative step toward digital health, empowering people to proactively manage their health journeys. The growing focus on energy efficiency and the adoption of eco-friendly use habits emphasize the significance of these apps. Research by Choi et al [7] and Pop et al [8] shed light on the essential role that energy-efficient software plays in prolonging the lifespan of devices and mitigating environmental impacts, heralding a significant shift in digital health practices. The fusion of wearable technologies with these apps further highlights the importance of designing with energy mindfulness at the forefront, ensuring that our pursuit of health does not lead to unsustainable energy use. Table 1 illustrates the relationship between the popularity of mHealth apps and their user review scores. Apps with the highest user satisfaction were selected in this study to be assessed for energy efficiency.

Table 1. Correlation between app popularity, where popularity is determined by the number of downloads.

App name	Downloads (in millions)	User review (out of 5)
Ate Food Journal	2	4.4
Calorie Counter	5	4.2
Lifesum	10	4.6
My Plate	8	4.4
MyFitnessPal	45	4.4
Noom	15	4.2
Ovia	3	4.0
PlateJoy	1	4.6
Spokin	4	4.2
Yummly	20	4.6

Problem Statement

Even though mHealth and nutrition apps have become increasingly popular, there is a dearth of research on how much energy they are consumed on Android devices and practical guidance on what can users do about it. Almasri and Gouveia [9] studied the gap in sustainable practice using Android apps and highlighted the need given their popularity and potential for energy-saving practice and given the global priority and commitment toward creating sustainable smartphones to achieve sustainable development goal 12.

Objective

The objective of this study is to assess the energy consumption of popular mHealth and nutrition apps and identify key areas where improvements can be made.

Literature Review

mHealth Apps and Energy Consumption

The widespread use of mHealth apps in our everyday routines has underscored the need to better understand energy consumption. Awais et al [10] examined the direct link between the complexity of these apps and their energy demands. Their findings indicate that apps with advanced features, such as real-time monitoring and personalized recommendations, can consume up to 30% more energy compared with simpler apps. Additionally, Sahar et al [11] provided an in-depth analysis of how unnoticed background activities, such as continuous data syncing and location tracking, play a significant role in draining smartphone batteries. Their study revealed that background activities could account for up to 40% of an app's total energy consumption, underscoring the importance of both developers and users to understand the app architecture and appreciate the influence it plays on energy use.

User Behavior and Energy Efficiency

Understanding energy efficiency warrants an understanding of user behaviors around app use. Personal relationships, belief in one's abilities as presented by Rahman et al [12] (self-efficacy), and the collective confidence in our shared power to effect change are key to embracing and consistently using mHealth technologies [13-15]. How individuals use mHealth apps has a significant impact on energy consumption patterns. Al Nidawi et al [16] showed that regular app use, including entering data and syncing, significantly increases energy consumption. Acer et al [17] highlighted how notifications, a common feature in mHealth apps, significantly boost energy use.

XSL•FO
RenderX

Strategies for Energy Optimization

Isuwa et al [18] showed that using adaptive brightness settings and energy-saving modes can extend the battery life of mobile devices by up to 20%. Furthermore, Benkhelifa et al [19] explored the potential of leveraging software-defined networking for energy optimization in mobile cloud computing, resulting in a decrease of up to 25% in energy use.

Technological Advancements and Energy Consumption

Emerging technologies play a nuanced role in the story of mHealth apps' energy consumption, presenting a mix of hurdles and breakthroughs. On the one hand, advancements in app development frameworks, as outlined by Kelényi et al [20], opened fresh opportunities for energy efficiency. On the other hand, the growing complexity of those apps, as pointed out by Porter [21], introduces significant obstacles to keeping energy use in check. The potential of artificial intelligence (AI) in optimizing energy consumption for sustainability has been highlighted in a recent article by Ericsson [22]. Their research indicates that AI features, while enhancing app functionality, can lead to a 25% increase in energy consumption if not optimized properly.

Cross-Platform Analysis of Energy Consumption

Khan et al [23] conducted a comparative analysis of power consumption in mobile devices to inform the development of energy-efficient mobile apps. Their study introduced a methodology for assessing and evaluating power use, providing valuable insights and guidelines for developers aiming to create more sustainable mobile apps. Ciman and Gaggi [24] analyzed smartphone energy consumption using different sensors, including either only app, or by using GPS, accelerometer, compass, camera, or microphone. They found that cross-platform frameworks significantly increase energy consumption compared with native apps. They suggested that power consumption should be considered when choosing between native implementation and using a framework or between different frameworks for mobile app development.

Energy Consumption Metrics and Measurement Techniques

Ergasheva et al [25] explored metrics of energy consumption to evaluate the energy efficiency of apps. They introduced metrics, such as energy-per-function, which quantifies the energy consumed for each app function, and energy-per-user interaction, which measures the energy used per user interaction, providing a more granular understanding of app energy consumption. Pathak et al [26] used advanced methods for tracking app energy consumption in real time, offering insights into the variables that drive energy use. They developed a real-time energy monitoring framework that captures detailed energy use data at the component level, enabling developers to identify energy hotspots within the app. This approach allows for more targeted energy optimization strategies, focusing on the most energy-intensive components and interactions.

Methods

Ethical Considerations

The approach we took was a quantitative analysis study measuring energy use in popular US-based health and nutrition apps. This study did not require ethics board approval as it involved the quantitative analysis of publicly available data related to the energy consumption of mHealth apps. No human subjects were directly involved, and no personal or sensitive data were collected during the study. This approach aligns with the institutional guidelines and adheres to regional and local policies regarding research involving nonhuman subjects, ensuring that all analyses remain within ethical boundaries as per the existing frameworks.

Selection of mHealth Apps

The selection of mHealth apps was based on Almasri and Gouveia's [27] criteria and insights from Kelényi et al [20].

Popularity and User Base

Apps with a vast number of downloads and positive feedback from users were selected.

Functional Complexity

Apps featuring a spectrum of functionalities were selected, from the simplest to the most complex, aiming to understand how different features influence energy use following the concept by Isuwa et al [18].

Energy Consumption Potential

Apps known or suspected to be high on energy use, including features such as continuous data syncing or GPS tracking, based on preliminary evaluations and what developers have documented (Benkhelifa et al [19]), were selected.

Measurement of Energy Consumption

Overview

Ergasheva et al [25] and Pathak et al [26] both introduced metrics such as energy-per-function and energy-per-interaction, offering detailed insights into app energy efficiency by measuring energy use for specific functions and user interactions. The process unfolded in 3 key steps as given below.

Baseline Measurement

We first set a baseline for energy consumption for each app when it was not in use, providing a benchmark for comparing energy use during more active scenarios.

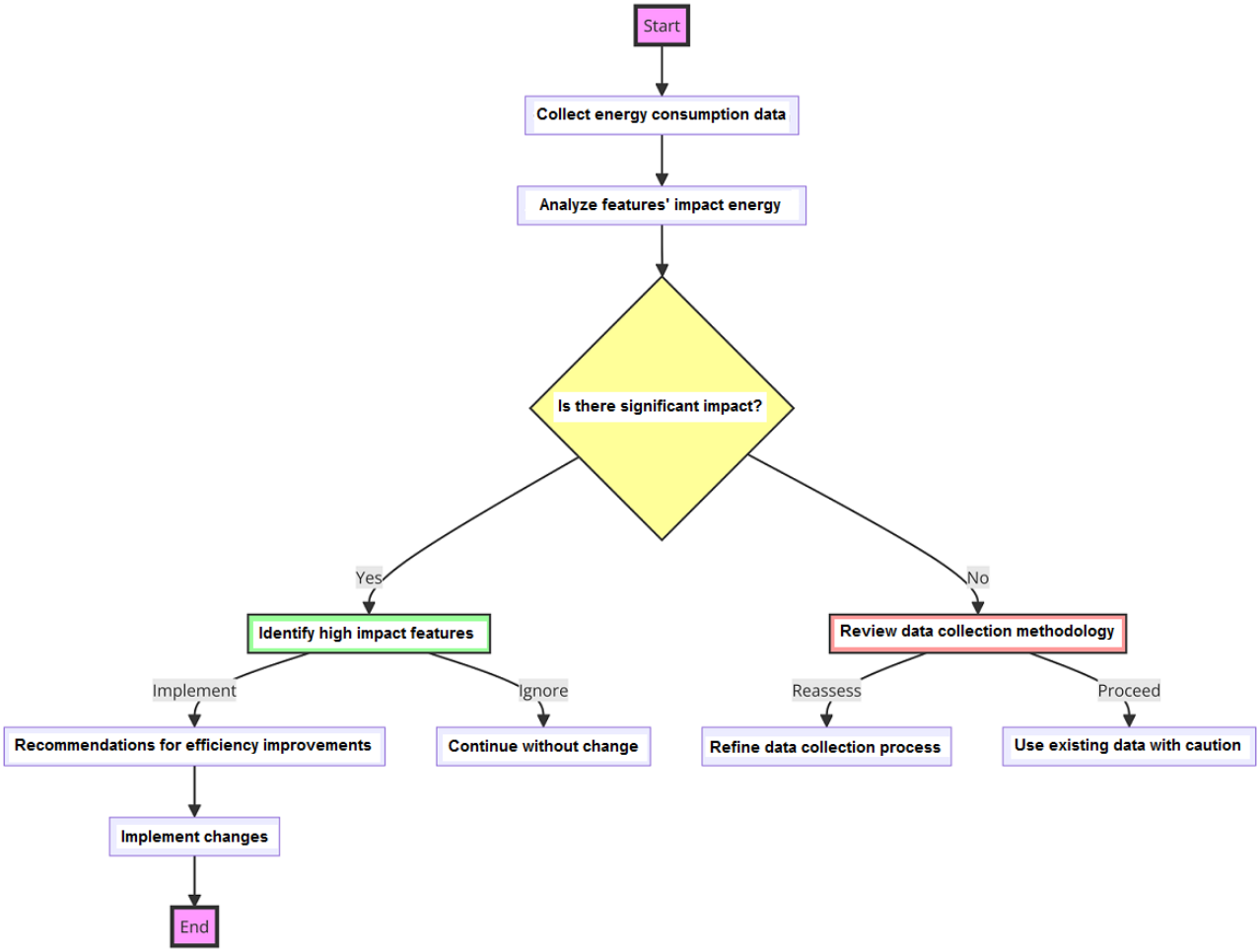
Feature-Specific Scenarios

We then measured energy use in scenarios that trigger specific features of the apps such as logging meals or syncing with wearable technology. This step was crucial for pinpointing and measuring the energy footprint of distinct functionalities within the apps. The flowchart depicted in Figure 1 outlines the sequential steps taken from the collection of energy consumption data to the identification of high-impact features and a review of the data collection methodology. Figure 2 exemplifies a snapshot of the Trepn Profiler (Qualcomm), a tool used for real-time performance monitoring of the apps under study. The

XSL•FO
RenderX

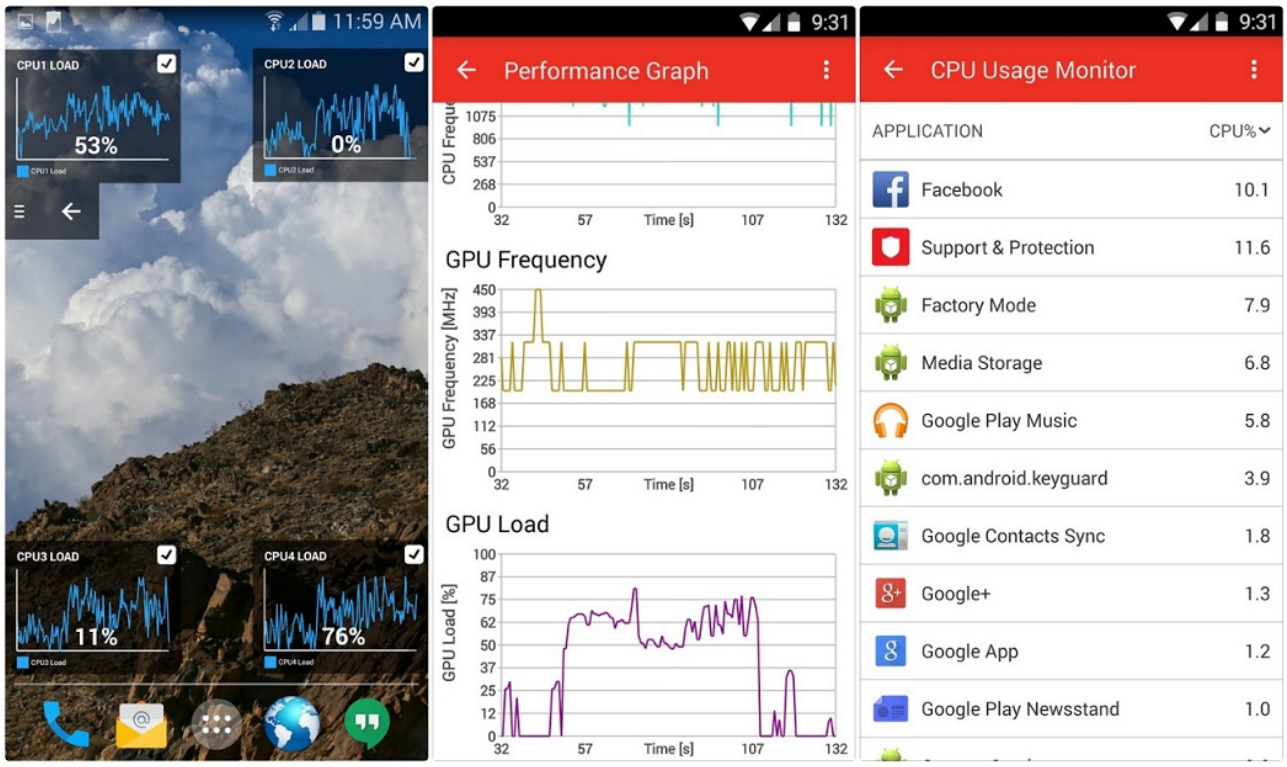
graphs depict central processing unit frequency and graphics processing unit load over a session, demonstrating how various app features and user interactions can influence energy consumption. Such detailed monitoring is indispensable for identifying high-energy-demand periods, thereby informing our strategies for app optimization.

Figure 1. Flowchart of the energy consumption analysis methodology.



XSL•FO
RenderX

Figure 2. Example of real-time performance monitoring using Trepn Profiler. CPU: central processing unit; GPU: graphics processing unit.



User Interaction Patterns

Finally, by simulating a range of real-life user interactions, from minimal to extensive use, we were able to depict how different use patterns impact the app's energy consumption.

Data Analysis

Overview

This study examined data on energy consumption and the potential of behavior change interventions to cut energy use, drawing inspiration from Internet of Things–enabled tactics designed to boost energy efficiency consistent with recent findings that underscore the effectiveness of behavioral strategies in curbing energy use across different contexts [28,29]. We conducted our statistical analysis using MATLAB software (MathWorks), and our analysis approach included descriptive statistics, comparative analysis, and regression analysis.

Descriptive Statistics

Energy consumption for each app across various scenarios was quantified in milliwatt-hour (mWh) using real-time energy monitoring tools. This analysis provided insights into energy consumption patterns and fluctuations.

Comparative Analysis

ANOVAs were used to compare energy data across different apps and scenarios, identifying significant differences attributable to app features or user interactions.

Regression Analysis

Building regression models enabled us to measure how certain factors, such as how often notifications pop up or GPS tracking is used, impact energy use. This analysis helps understand the layers of what drives energy consumption in mHealth apps.

Results

Energy Consumption Patterns

Table 2 maps out a comparative analysis of the 10 popular mHealth apps. It elucidates each app's market presence, user experience, and estimated energy consumption, laying a foundation for understanding the interplay between app features and energy efficiency. We found a notable range in how much energy these apps use, with some consuming up to 3 times more energy than their counterparts in similar conditions. This disparity stemmed from various factors, such as the complexity of the app's features, how efficiently it runs in the background, and how often and in what ways users interact with the app. In our analysis, we conducted a descriptive statistical examination to highlight the energy consumption patterns of the selected apps. Our findings reveal a variance in average energy use, with apps consuming between 150 mWh and 310 mWh under typical use scenarios. The SD in energy consumption underscores the variability, ranging from 15 mWh to 31 mWh, which is indicative of how user interactions and background processing contribute to energy expenditure. The minimum and maximum energy use values further allocate the range of energy efficiency among these apps, from 135 mWh to 341 mWh, reflecting the impact of app features and optimization on battery life.

XSL•FO
RenderX

Table 2. Detailed comparative analysis of top mobile health apps.

App name	Popularity (downloads)	User review (out of 5)	Average energy use (CPU[a])	Feature complexity	Integration with wearables	Notification frequency	Support for multiple diets
Yummly	>5M[b]	4.3	Low	Medium	No	Low	Yes
My Plate Calorie Counter	>5M	4	Medium	Medium	No	Medium	Yes
MyFitnessPal	>50M	4.6	High	High	Yes	Medium	No
Spokin	>50K[c]	3.9	Low	Low	No	Low	No
PlateJoy	>500K	4.2	Low	Low	No	Low	Yes
Ovia	>1M	4.1	Medium	Low	No	Medium	No
Lifesum	>10M	4.5	Medium	High	Yes	High	Yes
Noom	>10M	4.4	High	High	Yes	High	No
Calorie Counter	>10M	4.7	High	High	Yes	High	Yes
Ate Food Journal	>100K	4.2	Low	Low	No	Low	No

[a]CPU: central processing unit.

[b]M: million or more.

[c]K: hundred thousand or more.

Table 3 illustrates the comparative energy consumption patterns of 10 popular mHealth apps under various use scenarios. This visualization underscores the substantial disparities in energy use, driven by factors such as app feature complexity, background processing efficiency, and user interaction methods. It highlights the critical need for targeted energy optimization strategies to mitigate the significant energy demands of feature-rich apps. The apps that demanded the most energy were those packed with sophisticated features such as live syncing with wearable tech and ongoing background updates. On the flip side, the more straightforward apps that relied on manual inputs and had fewer background processes were much kinder to battery life.

Table 3. Comparative energy consumption patterns of 10 popular mobile health apps.

App name	Use scenario and energy consumption (milliwatt-hour)		
	Baseline	GPS use	High use
Ovia	1	4	22
Calorie Counter	3	5	18
My Plate Calorie Counter	5	8	16
Yummly	5	10	20
Lifesum	4	9	17
Noom	3	7	15
My Fitness Pal	2	5	23
Ate Food Journal	1	6	13
PlateJoy	2	8	11
Spokin	3	6	10

Additionally, our research highlights how the use of notifications and alerts plays a significant role in energy consumption. Apps that leaned heavily on notifications to keep users engaged were more likely to use more energy, primarily due to the frequent lighting up of screens and the data exchanged over network services. It was found that on average, using an mHealth app for an hour each day could drain approximately 15% to 20% of a smartphone's battery life, depending on the app's complexity and background activity. This observation points to the critical need for fine-tuning notification strategies, and finding a sweet spot that maintains user interest without unnecessarily draining the battery.

Impact of App Features on Energy Use

Our findings suggest that certain app functions are linked to the amount of energy they use. GPS tracking—used for recording outdoor meals or activities—along with frequent data synchronization and sophisticated graphical interfaces, emerged as the main factors driving up energy consumption. GPS tracking was particularly notable for its high energy use, relying heavily on constant location services and data exchange. The

XSL•FO
RenderX

research also brings to light how user behavior affects energy use, specifically how long and how often people use the apps. Apps designed to keep users engaged for longer periods, whether through fun gamification features or detailed dietary logging, were seen to consume more energy overall. This finding points to the need for thoughtfully crafting user engagement methods to avoid unnecessary energy consumption. Our findings are consistent with existing research. Choi et al [7] reaffirmed the considerable effect of screen brightness and network use on energy consumption. Our findings highlight the intense energy demands of certain features in mHealth apps, such as GPS tracking and frequent synchronization, areas not deeply researched by previous studies. While earlier research underscored the significance of hardware and system optimizations for lowering energy use, our research emphasizes the paramount role of optimizations at the app level where user-centered design and behavior will be critical. By concentrating on the architecture and capabilities of mHealth apps, developers have a profound opportunity to enhance the energy efficiency of their creations, as well as by involving both users and practitioners alike who can guide what features remain paramount for impactful and sustainable technology practice.

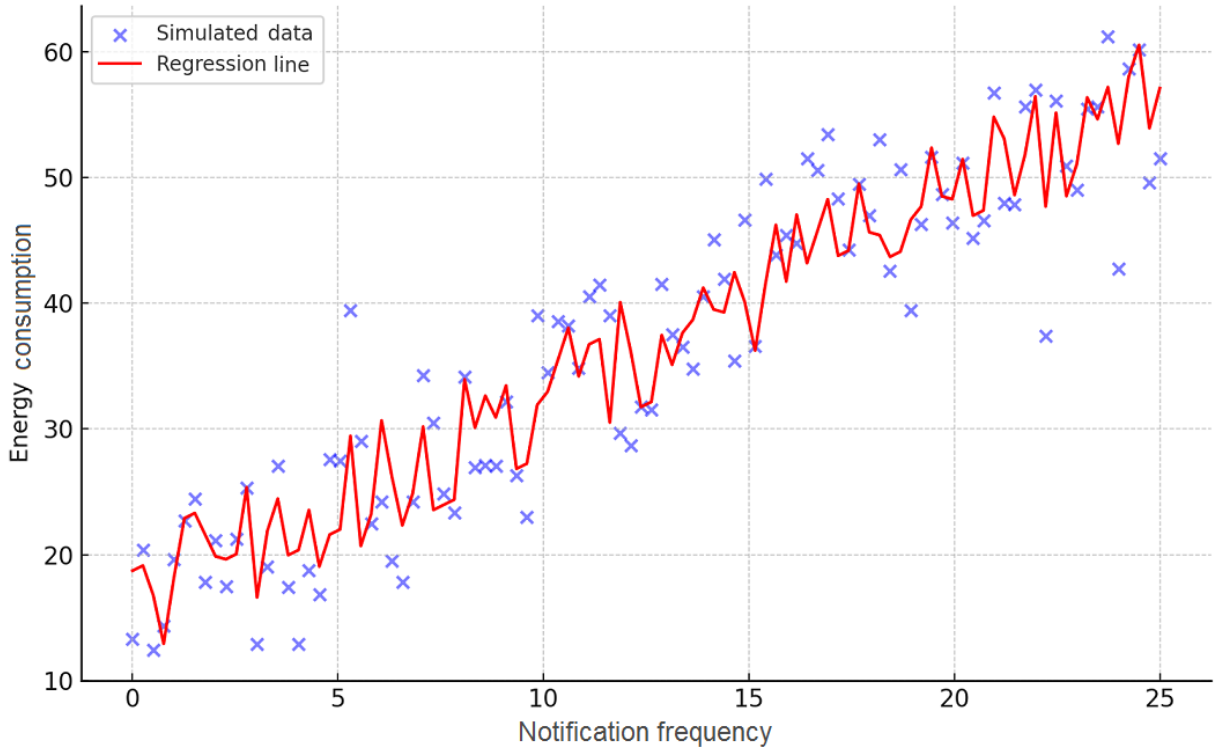
Insights From User Engagement and Energy Efficiency

Our regression analysis showed the relationship between user engagement, app features, and energy consumption. The analysis featured a key insight: while increased user engagement typically leads to higher energy consumption, strategic app design can mitigate this effect. Specifically, our findings highlight how certain app features, such as notification frequency, GPS use, and complexity level, influence the energy efficiency of mHealth apps. Assuming a linear relationship between these factors and energy consumption, our regression model is represented by the equation:

$$\text{Energy consumption} = \beta_0 + \beta_1 \times \text{notification frequency} + \beta_2 \times \text{GPS use} + \beta_3 \times \text{app complexity} + \varepsilon$$

where β_0 is the intercept, indicating the baseline energy consumption in the absence of the examined features. β_1, β_2, and β_3 are coefficients quantifying the impact of notification frequency, GPS use, and app complexity on energy consumption, respectively, and ε represents the error term, accounting for variability not explained by the model. Figure 3 clarifies this relationship, presenting a regression analysis that demonstrates the impact of notification frequency on energy consumption.

Figure 3. Regression analysis of notification frequency on energy consumption.



Our simulated analysis yielded the following equation:

$$\text{Energy consumption} = 9.55 + 1.62 \times \text{notification frequency}$$

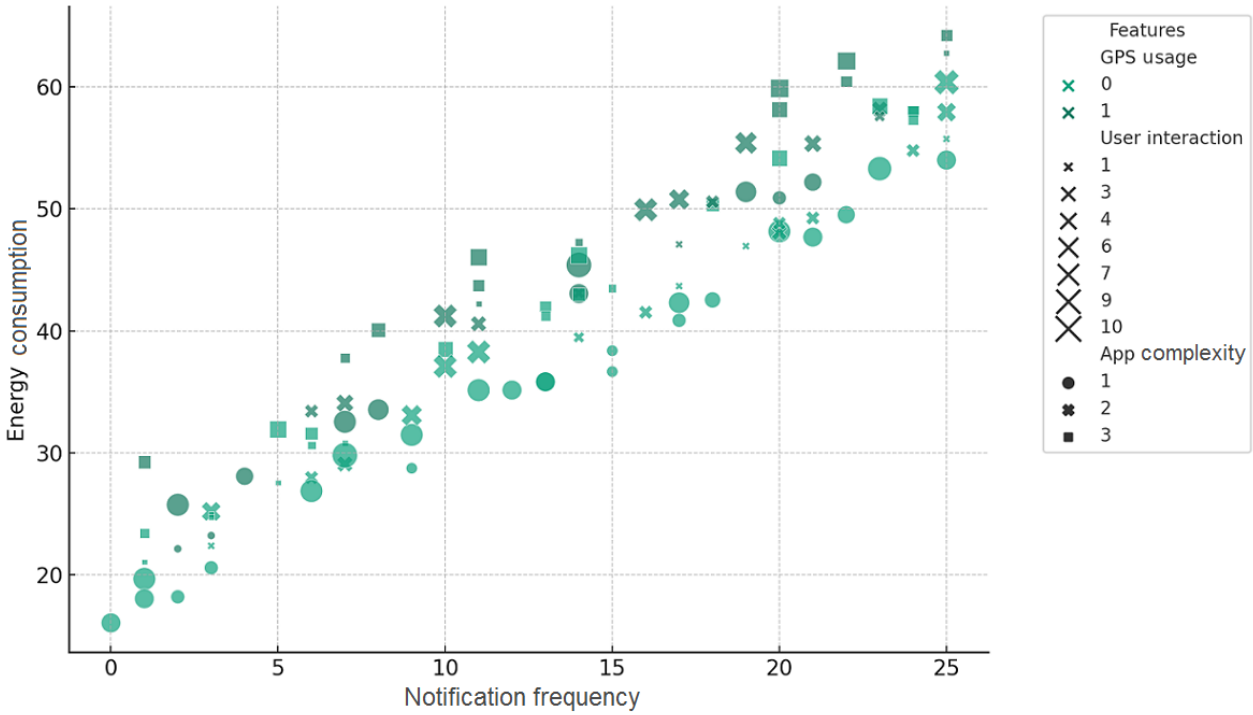
The coefficients derived from our analysis which provide insights into the relative influence of each feature on energy consumption start with the intercept (β_0=9.55) representing the baseline energy consumption. Then, each unit increase in notification frequency (β_1=1.62) corresponds to a 1.62-unit increase in energy consumption, emphasizing its significant role. Besides, GPS use (β_2=5.00) suggests that activating GPS functionality contributes an additional 5 units to energy consumption, highlighting the substantial energy demand of location services. Finally, the app complexity (β_3=3.00) shows that higher complexity levels increase energy consumption by 3 units, indicating the impact of advanced features and functionalities. The statistical significance of each coefficient

XSL•FO
RenderX

was evaluated through *P* values, confirming the strength of our findings. The *P* value for notification frequency is .01, indicating a highly significant relationship with energy consumption. The *P* value for GPS use is .05, suggesting an impact on energy consumption at the 5% level. The *P* value for app complexity is .03, demonstrating its significant effect on energy consumption.

For example, apps that adopt flexible synchronization schedules and energy-conscious notification strategies are more likely to succeed in keeping users engaged without a corresponding spike in energy use. This finding is also critical for developers aiming to refine the user experience while staying true to the principles of energy efficiency. Furthermore, we include an "integrated analysis approach" to examine the compounded effects of app features on energy consumption. This analysis builds upon our original regression model by openly considering the interactions between different app functionalities and their collective impact on energy use. To convey this concept, in Figure 4, we present an integrated analysis, contrasting the specific energy demands of app features against user interaction patterns, highlighting the potential for energy optimization. This visualization highlights the synergy between GPS use, notification frequency, app complexity, and their aggregate effect on energy use. Through this analysis, we aim to guide developers in identifying which combinations of features escalate energy demand and how thoughtful integration can mitigate such effects, fostering more energy-efficient app designs.

Figure 4. Integrated analysis of feature-specific energy consumption and user interaction patterns.



Discussion

Principal Findings

Finding the right equilibrium among app functionality, user satisfaction, and energy efficiency is critical for sustainable practice. Our research analysis enriches our understanding of the nuanced relationships within mHealth app use and highlights the broader consequences and opportunities for innovation in the digital health domain, underscoring the importance of a balanced approach to app development that honors both human and environmental considerations.

The diverse energy use among various mHealth apps, especially those dedicated to diet and meal tracking, reveals how an app is built and how users interact with it. Features, such as GPS tracking and constant data updates, significantly increase energy consumption, emphasizing the urgency for app creators to weave energy efficiency into the fabric of app development. User behaviors are also critical—including how often users interact with the app, respond to notifications, or use specific features. Behavior change modalities also need to be introduced to address user habits and smarter app configurations.

Our study findings bring forth the question of the need to identify modalities or to balance between incorporating features that boost user engagement and satisfaction and the essential task of reducing energy consumption. Such modalities are important to prolong battery life and lessen the ecological impact of mHealth app use.

Implications for Developers and Users

This study highlights for developers the crucial role of weaving energy efficiency into every stage of the app development cycle. This means going beyond just streamlining code and choosing low-power software development code libraries. It also means crafting app features and user interactions in ways that naturally lead to less energy use. Developers are urged to embrace smart algorithms that dynamically tweak app functions according to real-time energy use and battery status, ensuring the apps are as energy-efficient as possible without sacrificing the quality of the user experience.

XSL•FO
RenderX

On the user side, the research points out how a little awareness about how apps are set up and used can go a long way in reducing energy consumption. Users have the power to drive energy savings by tweaking their app settings, such as reducing how often apps search for new content or turning off unnecessary background activities. It is paramount to use apps that are designed from the ground up to be energy conscious with efficient battery life and for the purpose of encouraging a greener approach to leveraging digital health tools.

Broader Implications for Digital Health Technology

This research adds a valuable perspective to the conversation about making digital health technologies more sustainable, emphasizing the collective responsibility of consumption and production, including developers, users, health care professionals, and stakeholders, to put energy efficiency at the forefront. As mHealth apps play a more prominent role in enhancing health and nutrition outcomes and managing diseases, understanding and optimizing their energy use becomes essential for ensuring digital solutions can grow sustainably, and consumers and producers are both responsible for sustainable practice as well.

Moreover, the insights gathered here highlight the exciting possibilities of interdisciplinary studies that merge knowledge from software engineering, behavioral science, and environmental sustainability. This approach could lead to the creation of comprehensive guidelines and best practices for crafting mHealth apps that are not only effective but also energy efficient. By working together across fields, there is a tremendous opportunity to drive forward app innovations that serve the dual purpose of advancing health care while respecting our planet.

Conclusion

The mHealth apps within the mHealth sector consume energy, especially when app functionalities are governed by how we interact with these apps. It is a challenge for developers and users to find the right mix of features that drive engagement and health and nutrition benefits while also becoming cognizant of reducing energy use.

For developers, this means weaving energy efficiency more deeply into the fabric of app creation, from concept through to coding. This can be done by embracing flexible technologies and applying forward-thinking design philosophies that marry efficacy with energy savings. For users, it is about becoming more aware of how the choices they make in app settings and their daily use can affect energy consumption, moving toward a more conscious and deliberate use of these digital tools.

Acknowledgments

The authors would like to acknowledge the support of Prince Sultan University for paying the article processing charges of this publication and for their support.

Conflicts of Interest

None declared.

References

1. Bhaskar MS, Almakhles D, Padmanaban S, Ionel DM, Blaabjerg F, He J, et al. Investigation of a transistor clamped T-type multilevel H-bridge inverter with inverted double reference single carrier PWM technique for renewable energy applications. IEEE Access 2020;8:161787-161804 [FREE Full text] [doi: 10.1109/access.2020.3020625]
2. Muthanna MSA, Muthanna A, Nguyen TN, Alshahrani A, Abd El-Latif AA. Towards optimal positioning and energy-efficient UAV path scheduling in IoT applications. Comput Commun 2022;191:145-160 [FREE Full text] [doi: 10.1016/j.comcom.2022.04.029]
3. Ashfaq T, Khalid MI, Ali G, Affendi ME, Iqbal J, Hussain S, et al. An efficient and secure energy trading approach with machine learning technique and consortium blockchain. Sensors (Basel) 2022;22(19):7263 [FREE Full text] [doi: 10.3390/s22197263] [Medline: 36236363]
4. Chandani A, Wagholikar S, Prakash O. A bibliometric analysis of green computing. In: Xie J, Kaiser MS, Rathore VS, editors. Information and Communication Technology for Competitive Strategies (ICTCS 2021): Intelligent Strategies for ICT. Springer: Springer Nature; 2023:547-557.
5. Switzer J, Siu E, Ramesh S, Hu R, Zadorian E, Kastner R. Renée: new life for old phones. IEEE Embedded Syst Lett 2022;14(3):135-138 [FREE Full text] [doi: 10.1109/les.2022.3147409]
6. Muchinenyika SH, Muyingi HNN. An analysis of screen resolution effects on battery endurance: a case of smartphones. 2022 Presented at: 2022 IEEE 9th International Conference on Sciences of Electronics, Technologies of Information and Telecommunications (SETIT); May 28-30, 2022; Hammamet, Tunisia p. 356-360 URL: https://ieeexplore.ieee.org/abstract/document/9875592?casa_token=l9Pq-gdVmZ0AAAAA:nUzSENDhR1EmpEzMUPBuJr7lxSsSvP9_5HLUjrLJRkRhhh3tEPLshJNFOtJBw1A_IYFNpRVg [doi: 10.1109/setit54465.2022.9875592]
7. Choi Y, Park S, Jeon S, Ha R, Cha H. Optimizing energy consumption of mobile games. IEEE Trans Mob Comput 2022;21(10):3744-3756 [FREE Full text] [doi: 10.1109/tmc.2021.3058381]

XSL•FO
RenderX

8. Pop RA, Dabija DC, Pelău C, Dinu V. Usage intentions, attitudes, and behaviors towards energy-efficient applications during the COVID-19 pandemic. J Bus Econ Manag 2022;23(3):668-689 [FREE Full text] [doi: 10.3846/jbem.2022.16959]
9. Almasri AM, Gouveia LB. The evolution of energy requirements of smartphones based on user behaviour and implications of the COVID-19 era. Int J Electr Comput Eng 2021;11(3):2423-2431 [FREE Full text] [doi: 10.11591/ijece.v11i3.pp2423-2431]
10. Awais Q, Adeel M, Jawad H, Adnan K. Evaluating the impact of design pattern usage on energy consumption of applications for mobile platform. Appl Comput Syst 2021 Jun 04;26(1):1-11. [doi: 10.2478/acss-2021-0001]
11. Sahar H, Bangash AA, Beg MO. Towards energy aware object-oriented development of android applications. Sustain Comput Inform Syst 2019;21:28-46 [FREE Full text] [doi: 10.1016/j.suscom.2018.10.005]
12. Rahman MS, Ko M, Warren J, Carpenter D. Healthcare technology self-efficacy (HTSE) and its influence on individual attitude: an empirical study. Comput Human Behav 2016;58:12-24 [FREE Full text] [doi: 10.1016/j.chb.2015.12.016]
13. Sanakulov N, Karjaluoto H. Consumer adoption of mobile technologies: a literature review. Int J Mob Commun 2015;13(3):244-275 [FREE Full text] [doi: 10.1504/ijmc.2015.069120]
14. Rahimi N, Jetter A. Explaining health technology adoption: past, present, future. 2015 Presented at: 2015 Portland International Conference on Management of Engineering and Technology (PICMET); August 02-06, 2015; Portland, OR, USA p. 2465-2495. [doi: 10.1109/picmet.2015.7273030]
15. Goddard RD, Salloum SJ. Collective efficacy beliefs, organizational excellence, and leadership. In: Spreitzer GM, Cameron KS, editors. The Oxford Handbook of Positive Organizational Scholarship. Oxford: Oxford University Press; 2012:642-650.
16. Al Nidawi HSA, Koh TW, Dawood K, Khaleel A. Energy consumption patterns of mobile applications in android platform: a systematic literature review. J Theor Appl Inf Technol 2017;95(24):6776-6787 [FREE Full text]
17. Acer UG, Mashhadi A, Forlivesi C, Kawsar F. Energy efficient scheduling for mobile push notifications. EAI Endorsed Trans Energy Web 2015;2(5):100-109. [doi: 10.4108/eai.22-7-2015.2260067]
18. Isuwa S, Amos D, Singh AK, Al-Hashimi BM, Merrett GV. Maximising mobile user experience through self-adaptive content-and ambient-aware display brightness scaling. J Syst Architect 2023;145:103023 [FREE Full text] [doi: 10.1016/j.sysarc.2023.103023]
19. Benkhelifa E, Welsh T, Tawalbeh L, Jararweh Y, Al-Ayyoub M. Leveraging software-defined-networking for energy optimisation in mobile-cloud-computing. Procedia Comput Sci 2016;94:479-484 [FREE Full text] [doi: 10.1016/j.procs.2016.08.074]
20. Kelényi I, Nurminen JK, Siekkinen M, Lengyel L. Supporting energy-efficient mobile application development with model-driven code generation. In: van Do T, Le Thi HA, Nguyen NT, editors. Advanced Computational Methods for Knowledge Engineering: Proceedings of the 2nd International Conference on Computer Science, Applied Mathematics and Applications (ICCSAMA 2014). Cham: Springer International Publishing; 2014:143-156.
21. Porter C. Energy efficiency in high-performance computing: balancing speed and sustainability. NVIDIA Technical Blog. 2023. URL: https://developer.nvidia.com/blog/energy-efficiency-in-high-performance-computing-balancing-speed-and-sustainability/ [accessed 2024-02-29]
22. Ericsson. Intelligent sustainability: the role of AI in energy consumption, management and new revenues. Ericsson Blog. 2023. URL: https://www.ericsson.com/en/blog/2023/6/intelligent-sustainability-ai-energy-consumption [accessed 2024-02-25]
23. Khan MU, Abbas S, Lee SUJ, Abbas A. Measuring power consumption in mobile devices for energy sustainable app development: a comparative study and challenges. Sustain Comput Inform Syst 2021;31:100589 [FREE Full text]
24. Ciman M, Gaggi O. Evaluating impact of cross-platform frameworks in energy consumption of mobile applications. 2014 Presented at: 10th International Conference on Web Information Systems and Technologies (WEBIST-2014); 3-5 April, 2014; Barcelona, Spain URL: https://www.scitepress.org/Papers/2014/48576/48576.pdf [doi: 10.5220/0004857604230431]
25. Ergasheva S, Khomyakov I, Kruglov A, Succil G. Metrics of energy consumption in software systems: a systematic literature review. IOP Conf Ser Earth Environ Sci 2020;431(1):012051 [FREE Full text] [doi: 10.1088/1755-1315/431/1/012051]
26. Pathak A, Jindal A, Hu YC, Midkiff SP. What is keeping my phone awake?: characterizing and detecting no-sleep energy bugs in smartphone apps. 2012 Presented at: MobiSys'12: The 10th International Conference on Mobile Systems, Applications, and Services; June 25-29, 2012; Low Wood Bay Lake District UK p. 267-280. [doi: 10.1145/2307636.2307661]
27. Almasri AM, Gouveia LB. Star-rating evaluation model for rating the energy-efficiency level of android google play apps. Int J Electr Comput Eng 2021;11(2):1599-1612 [FREE Full text] [doi: 10.11591/ijece.v11i2.pp1599-1612]
28. Ramallo-González AP, Bardaki C, Kotsopoulos D, Tomat V, Vidal AG, Ruiz PJF, et al. Reducing energy consumption in the workplace via IoT-allowed behavioural change interventions. Buildings 2022;12(6):708 [FREE Full text] [doi: 10.3390/buildings12060708]
29. Varvello M, Livshits B. Shedding (Some) light on mobile browsers energy consumption. Traffic Monitoring and Analysis. 2021. URL: https://tma.ifip.org/2021/wp-content/uploads/sites/10/2021/08/tma2021-paper18.pdf [accessed 2024-04-30]

Abbreviations

AI: artificial intelligence
mHealth: mobile health

XSL•FO
RenderX

mWh: milliwatt-hour

Edited by G Eysenbach, A Kushniruk, E Borycki; submitted 12.03.24; peer-reviewed by M Ayiad, Y Boukili, G Nogueira, N Russ; accepted 07.04.24; published 10.05.24.

Please cite as:
Almasri A, El-Kour TY, Silva L, Abdulfattah Y
Evaluating the Energy Efficiency of Popular US Smartphone Health Care Apps: Comparative Analysis Study Toward Sustainable Health and Nutrition Apps Practices
JMIR Hum Factors 2024;11:e58311
URL: https://humanfactors.jmir.org/2024/1/e58311
doi:10.2196/58311
PMID:38729624

©Abdullah Almasri, Tatyana Y El-Kour, Liliana Silva, Yousef Abdulfattah. Originally published in JMIR Human Factors (https://humanfactors.jmir.org), 10.05.2024. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on https://humanfactors.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Accessibility, Relevance, and Impact of a Symptom Monitoring Tool for Home Hospice Care: Theory Elaboration and Qualitative Assessment

Karla T Washington[1], PhD; Debra Parker Oliver[1,2], PhD; Allison K Donehower[1,3], MS; Patrick White[1,4], MD, PhD; Jacquelyn J Benson[1], PhD; Patrick G Lyons[5,6], MSc, MD; George Demiris[7,8], PhD

[1]Department of Medicine, School of Medicine, Washington University in St. Louis, St. Louis, MO, United States

[2]Goldfarb School of Nursing, Barnes-Jewish College, St. Louis, MO, United States

[3]Department of Human Development and Family Science, College of Education and Human Development, University of Missouri, Columbia, MO, United States

[4]BJC Hospice, St. Louis, MO, United States

[5]Department of Medicine, School of Medicine, Oregon Health & Science University, Portland, OR, United States

[6]Department of Medical Informatics and Clinical Epidemiology, School of Medicine, Oregon Health & Science University, Portland, OR, United States

[7]Department of Biostatistics, Epidemiology, & Informatics, School of Medicine, University of Pennsylvania, Philadelphia, PA, United States

[8]Department of Biobehavioral Health Sciences, School of Nursing, University of Pennsylvania, Philadelphia, PA, United States

Corresponding Author:
Karla T Washington, PhD
Department of Medicine
School of Medicine
Washington University in St. Louis
Mid Campus Center
4590 Children's Place, Mailstop 90-29-931
St. Louis, MO, 63110
United States
Phone: 1 (314) 285 0905
Email: kwashington@wustl.edu

Abstract

Background: Early users found Engagement and Visualization to Improve Symptoms in Oncology Care (ENVISION), a web-based application designed to improve home management of hospice patients' symptoms and support patients' and family caregivers' well-being, to be generally useful and easy to use. However, they also raised concerns about potential challenges users with limited technological proficiency might experience.

Objective: We sought to concurrently accomplish two interrelated study aims: (1) to develop a conceptual framework of digital inclusivity for health information systems and (2) to apply the framework in evaluating the digital inclusivity of the ENVISION application.

Methods: We engaged ENVISION users (N=34) in a qualitative study in which data were collected via direct observation, think-aloud techniques, and responses to open-ended queries. Data were analyzed via theory elaboration and basic qualitative description.

Results: Accessibility, relevance, and impact were identified as 3 essential considerations in evaluating a health system's digital inclusivity. Study findings generally supported ENVISION's digital inclusivity, particularly concerning its perceived relevance to the work of family caregivers and hospice clinicians and its potentially positive impact on symptom management and quality of life. Limitations to ENVISION's digital inclusivity centered around issues of accessibility, particularly operability among individuals with limited technological knowledge and skills.

Conclusions: The Accessibility, Relevance, and Impact conceptual framework of digital inclusivity for health information systems can help identify opportunities to strengthen the digital inclusivity of tools, such as ENVISION, intended for use by a broad and diverse range of users.

XSL•FO
RenderX

(JMIR Hum Factors 2024;11:e51789) doi:10.2196/51789

KEYWORDS

caregivers; home care services; hospice care; signs and symptoms; technology; mobile phone

Introduction

Background

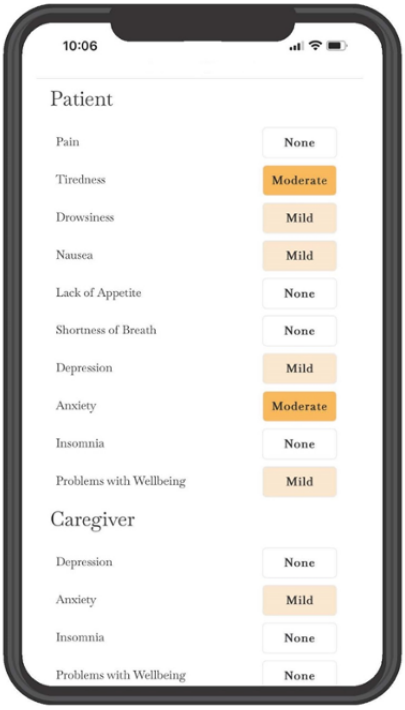
Hospice is a health care delivery model and a philosophy of care focused on reducing pain and promoting quality of life among patients who are terminally ill and their families [1]. In the United States, hospice care is most often provided in patients' homes [2]. While more intensive staffing is available during acute medical crises, routine home hospice care consists of only periodic visits from nurses, nursing aides, social workers, chaplains, and others operating under the direction of a hospice physician [3]. A total of 3- to 4-hour-long weekly home visits may be typical for an established patient. Thus, responsibility for the overwhelming majority of home hospice care falls to patients' family members and friends (referred to as *family caregivers*), who are typically unpaid and often lack formal health care training [4-6].

Hospice family caregivers are commonly tasked with in-home management of patients' symptoms, including pain, shortness of breath, anxiety, and fatigue. Recent population-based research indicates that >78% of family caregivers who assist with symptom management in the last month of a patient's life report difficulty doing so [7]. These findings are consistent with those of numerous other studies highlighting the reality that symptom management challenges are a significant source of stress for many hospice family caregivers [8-11]. These challenges, coupled with lack of a standardized processes for real-time symptom reporting and monitoring in home hospice care, commonly result in suboptimal home management of patients' symptoms [12].

Engagement and Visualization to Improve Symptoms in Oncology Care

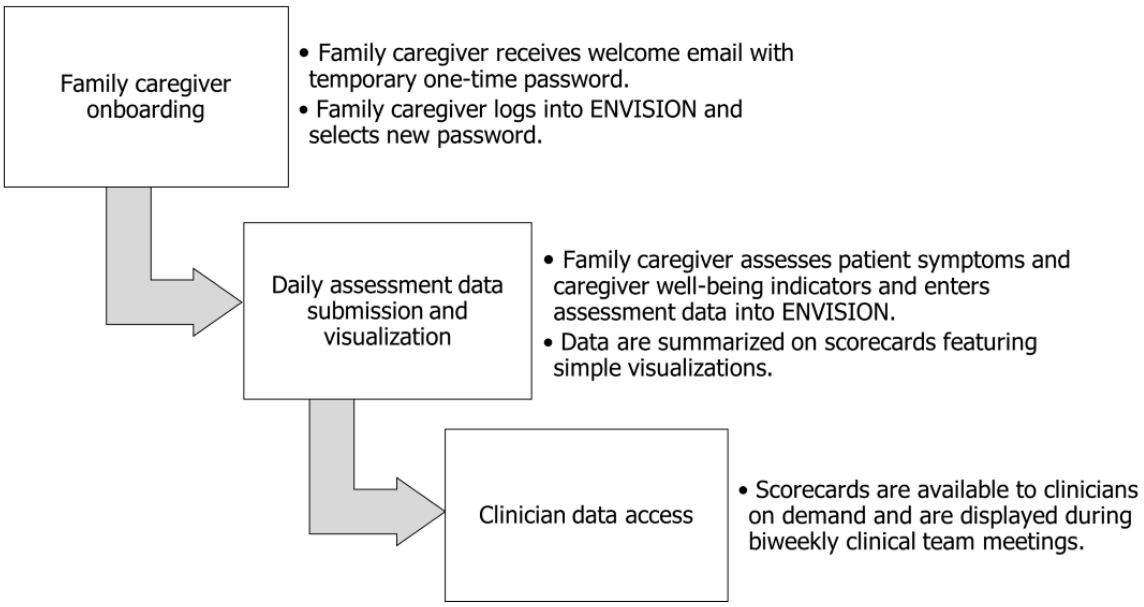
Engagement and Visualization to Improve Symptoms in Oncology Care (ENVISION) is a secure, web-based application designed to improve home management of hospice patients' symptoms and support patients' and family caregivers' well-being by improving the exchange of information between family caregivers and hospice clinicians [13]. It uses daily symptom and well-being data entered on the internet by family caregivers to create simple visualizations summarized in a patient and caregiver scorecard (Figure 1), allowing hospice clinicians to quickly identify areas of concern. These scorecards are displayed during biweekly hospice interdisciplinary team meetings and are available on demand to hospice clinicians outside of regularly scheduled meetings. A workflow diagram illustrating ENVISION's use is provided in Figure 2. Optional views, including longitudinal graphs of individual or combined symptoms, are also available to clinician users. Although ENVISION was initially developed specifically for advanced cancer care, its use has been expanded to include care for hospice-eligible individuals experiencing any life-limiting illness.

Figure 1. Mobile version of the application Engagement and Visualization to Improve Symptoms in Oncology Care (ENVISION), showing a sample patient and caregiver scorecard.



XSL•FO
RenderX

Figure 2. Engagement and Visualization to Improve Symptoms in Oncology Care (ENVISION) workflow.



Digital Inclusivity

ENVISION was created over several years with significant involvement of hospice family caregivers, clinicians, and administrators. While early research broadly supported ENVISION's usefulness and ease of use, it also raised concerns about potential barriers to use that might be experienced by family caregivers with limited technological skills or resources [13]. These concerns echo those voiced as part of an ongoing discussion in health care regarding digital inclusivity [14], broadly defined as "the ability of individuals and groups to access and use information and communication technologies [15]."

Digital inclusivity, particularly as it pertains to digital health technologies, is a salient concern on multiple social levels [16]. Individually, users vary with regard to their level of digital literacy and their ability to personally take advantage of available technological resources [17]. For example, an individual may struggle to discern differences between trustworthy and untrustworthy sources of health information, may have functional limitations (for example, vision or hearing impairment), or may be unable to afford home internet access. Similarly, families and other social groups differ in their degree of collective technological resources, a reality evident when less technologically equipped individuals benefit from the digital knowledge and skills of family members and friends. A family member assisting a patient in accessing their health care portal would be one example [18]. Communities can also be considered more or less digitally inclusive based on the adequacy of the infrastructure (such as home broadband connectivity or public Wi-Fi networks) available to meet residents' technological needs [15].

Study Aims

Our initial aim in conducting the study described herein was to better understand early ENVISION users' concerns regarding the application's digital inclusivity. However, in planning our study, we struggled to identify an existing conceptual framework to guide our research, given our plan to explore digital inclusivity as a quality of an individual application (rather than, for example, a community). Thus, we added a second study aim: to engage ENVISION users in a process of theory elaboration, resulting in a conceptual framework of digital inclusivity for health information systems. In this way, we sought to inform future ENVISION enhancements while contributing to the broader emerging science of digital inclusivity in health care. Thus, our finalized study aims were as follows: (1) to develop a conceptual framework of digital inclusivity for health information systems and (2) to apply the framework in evaluating the digital inclusivity of the ENVISION application.

Methods

Setting, Participants, and Recruitment

As part of ENVISION's ongoing, iterative, user-centered design [19], we recruited hospice family caregivers and clinicians (nurses, physicians, social workers, and chaplains) to participate in a qualitative research study [20]. We partnered with the university's Institute of Clinical and Translational Sciences' Recruitment Enhancement Core to recruit hospice family caregivers via flyers, targeted email blasts, social media posts, and listing of the study on a public-facing website. Family caregivers were eligible for inclusion if they were aged ≥18, able to speak and read English, and current or former (within the prior year) family caregivers of a patient receiving services from a Medicare-certified US hospice agency. We recruited hospice clinicians via social media posts and email blasts from professional hospice organizations, targeted emails to prior research partners, and presentation of the study opportunity at scheduled meetings of hospice agencies with which we had established partnerships. Hospice clinicians were eligible for study inclusion if they were aged ≥18, able to speak and read English, and currently employed or affiliated with a Medicare-certified US hospice agency.

XSL•FO
RenderX

Data Collection

All consenting participants met online individually with a research team member for approximately 30 to 45 minutes via a university-managed, Health Insurance Portability and Accountability Act–compliant Zoom (Zoom Video Communications) account [21]. At the start of each call, the researcher provided assistance in using Zoom's screen share feature, which the researcher later used to observe the participants completing a series of structured tasks in the ENVISION application. In addition, the researcher provided instruction in the think-aloud technique [22,23], explaining that the participants would be asked to verbalize their thinking as they navigated the application and completed specific tasks. The researcher also informed the participants that they would be asked to answer a series of open-ended questions about their perceptions of the application and its potential use in hospice care. Finally, the researcher confirmed the participants' understanding and began recording the session with the participants' knowledge and permission.

Structured Tasks

During the recorded Zoom session, family caregivers were sent a personalized email with a brief welcome message, a link to the ENVISION website, and a temporary one-time password (an automatically generated alphanumeric string of characters). As their first observed task, family caregivers were asked to navigate to the ENVISION website, enter the site using their email address and one-time password, and choose a new password. They were then asked to recall a typical caregiving day and enter corresponding symptom and well-being data for the patient and themselves into the ENVISION application. Next, they were asked to navigate to the patient and caregiver scorecard, which summarized the data they had just entered, and answer questions that required them to interpret simple data visualizations (labeled rectangles filled with different shades of orange ranging from none or white to dark or bright orange to reflect greater symptom intensity). Finally, they were asked to exit the application. After completing these structured tasks, they were asked a series of open-ended questions, including, for example, "What made it easy to use ENVISION?" "What made it challenging?" and "Which symptom(s) would be most important to communicate to the hospice team? Why?"

Hospice clinicians were also observed navigating to the ENVISION website, entering the site using their email address and temporary password, and selecting a new password. Because the researcher had entered them into the system as a clinician when generating their welcome email, clinicians were taken to a screen that included a list of fictitious patients' names and medical record numbers upon logging into the system. When they reached this screen, clinicians were asked to navigate to a specific patient's information (this required them to locate and click on the patient's name, but they were not given these specific instructions). Clicking on the patient's name took them to a screen that included a daily patient and caregiver scorecard that featured visualizations similar to those shown to family caregivers. This page also included a simple line graph that allowed clinician users to view the intensity of one or more symptoms or well-being indicators over time by clicking a box next to the appropriate symptoms or indicators (users were not provided with these specific instructions). While on this page, clinicians were asked questions that required them to interpret the colored boxes on the patient and caregiver scorecard; select and deselect specific symptoms on the longitudinal graph; and interpret trends, including symptom co-occurrence over time, shown via the graphed data. Finally, they were asked to exit the system and answer a series of open-ended questions, including, for example, "How, if at all, would having [information provided via ENVISION] affect how you do your job?" and "Describe how you would access ENVISION. For example, would you use a desktop computer, tablet, or smartphone? Would you use the application from the hospice agency office, patients' homes, or elsewhere?"

Data Preparation

In preparation for data analysis, we contracted with a third-party service to transcribe audio files of participants' recorded Zoom sessions verbatim. We then imported the resulting transcripts into NVivo qualitative analysis software (Lumivero). Complete copies of all audio and video files of participants' recorded Zoom sessions and corresponding field notes were stored in a secure Box folder made available to all institutional review board–approved research team members throughout data analysis.

Data Analysis

Our analysis was broadly informed by the work of organizational management researchers Fisher and Aguinis [24], who described a process they referred to as theory elaboration, defined as "the process of conceptualizing and executing empirical research using preexisting conceptual ideas or a preliminary model as a basis for developing new theoretical insights by contrasting, specifying, or structuring theoretical constructs and relations to account for and explain empirical observations." As part of this process, we engaged in vertical contrasting, which entailed adapting an existing conceptual framework (described in detail in the next paragraph) developed for one level of analysis to examine a phenomenon at another level. In doing so, we sought to determine which aspects of the framework functioned similarly on both levels of analysis and which functioned differently. We also engaged in construct specification, seeking to refine the constructs articulated in the original framework and to introduce new constructs when the existing constructs failed to capture important aspects of the phenomenon under investigation (in our case, ENVISION's digital inclusivity). At times, this involved construct splitting, a process whereby we split existing constructs into more specific dimensions if more conceptual specificity was needed to capture important aspects of ENVISION's digital inclusivity. Finally, we engaged in structuring, or identifying relationships between and among constructs, remaining open to new relation structures.

To accomplish these analytic activities (ie, contrasting, specifying, and structuring), 2 researchers (KTW and AKD) first reviewed all study transcripts, video files, and field notes. They then met to develop an initial codebook based on the elements of an existing framework: *Building Digital Communities: A Framework for Action* [15], created by the Institute of Museum and Library Services, the University of

XSL•FO
RenderX

Washington Technology and Social Change Group, and the International City or County Management Association. As its name suggests, this framework was created to promote digital inclusivity at the community level. Consistent with this purpose, it articulated 13 principles for community-wide digital inclusivity, including access principles (which addressed community infrastructure needs), adoption principles (which focused on community members' facilitators and barriers to use), and application principles (which specified areas where deployment of digital technologies would be likely to enhance community members' lives). The original framework's principles and corresponding definitions are provided in Multimedia Appendix 1.

We originally envisioned development of the initial codebook as a relatively straightforward process in which most, if not all, constructs articulated in the original framework would be initially retained and then refined in subsequent analytic steps. However, it soon became apparent that some of the original principles had limited applicability in the context of an individual application and should, therefore, be de-emphasized or excluded in the early stages of our analysis. For example, the application principles outlined in the original framework identified specific community sectors, such as education and public safety, where the deployment of technologies was deemed likely to benefit community well-being. However, digital health tools are, by definition, intended for deployment in health care (and, in the case of ENVISION, more specifically in hospice care). Thus, we omitted them from the initial codebook, feeling confident they would neither enrich our understanding of ENVISION's digital inclusivity nor ultimately represent constructs comprising our adapted conceptual framework.

After completing the initial codebook, KTW and AKD independently coded approximately 15% of the study transcripts, consulting video recordings and field notes as needed for context or clarification of transcribed data. KTW and AKD then met to make substantive modifications to the codebook to enhance its goodness of fit with the data. Examples of changes made at this stage included specifying that *affordability* referred to ENVISION's initial and ongoing costs and should, thus, be relabeled as *affordability and sustainability* (construct specification) and dividing *design for inclusion* into *perceivability*, *operability*, and *comprehensibility* (construct splitting), as the available data suggested that these were conceptually meaningful distinctions. We then used the modified codebook to code the entire data set, meeting afterward to compare individual coding decisions (resolving discrepancies via discussion and arriving at consensus), finalize our code definitions, and group related codes into broader categories that comprised our resulting conceptual framework and shed light on ENVISION's strengths and limitations with regard to digital inclusivity.

Ethical Considerations

All research activities were reviewed and approved by the Washington University in St Louis Institutional Review Board (#202105172).

Individuals interested in study participation were provided with contact information for our study coordinator, who screened potential participants for eligibility, obtained verbal informed consent for those interested in participating, and coordinated all subsequent research activities including participant payments of US $40 sent via check to the mailing address of the participants' choice.

Results

Overview

A total of 34 individuals participated in our qualitative research study, enabling the concurrent achievement of 2 interrelated study aims: (1) to develop a conceptual framework of digital inclusivity for health information systems and (2) to evaluate the digital inclusivity of the ENVISION application (participant characteristics are summarized in Table 1). In the following sections, we present our study findings, beginning with a brief overview of our conceptual framework and its essential elements. We then provide an in-depth description of the framework, illustrating its specific constructs with examples from our evaluation of ENVISION's digital inclusivity.

XSL•FO
RenderX

Table 1. Summary of participant characteristics (N=34).

Characteristic	Family caregivers (n=10), n (%)	Hospice clinicians (n=24), n (%)
Age range (y)		
18-29	1 (10)	2 (8)
30-39	1 (10)	5 (21)
40-49	1 (10)	6 (25)
50-59	2 (20)	8 (33)
60-69	3 (30)	3 (13)
≥70	2 (20)	0 (0)
Gender		
Man	1 (10)	5 (21)
Woman	9 (90)	19 (79)
Race		
Black	3 (30)	0 (0)
White	7 (70)	24 (100)
Ethnicity		
Hispanic	0 (0)	1 (4)
Non-Hispanic	10 (100)	23 (96)
Relationship to patient		
Spouse or partner	2 (20)	N/A[a]
Adult child	5 (50)	N/A
Other	3 (30)	N/A
Highest formal education		
Some college or trade school	2 (20)	N/A
Associate's degree	2 (20)	N/A
Bachelor's degree	3 (30)	N/A
Graduate or professional degree	3 (30)	N/A
Profession		
Chaplain	N/A	5 (21)
Nurse	N/A	7 (29)
Other	N/A	1 (4)
Physician	N/A	3 (13)
Social worker	N/A	8 (33)
Professional experience (y)		
0-5	N/A	6 (25)
6-10	N/A	5 (21)
11-15	N/A	4 (17)
16-20	N/A	2 (8)
21-25	N/A	2 (8)
>25	N/A	5 (21)

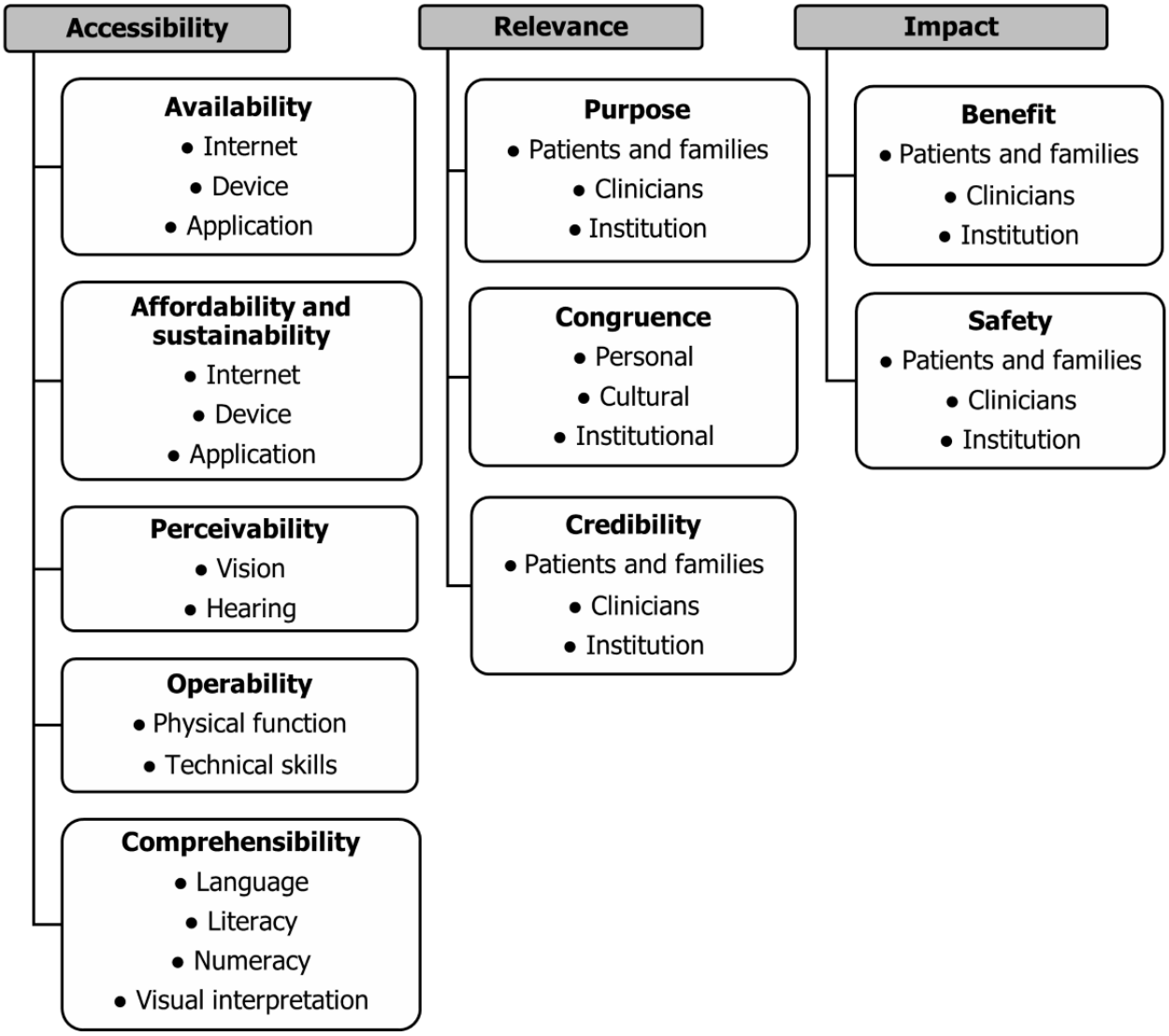
[a]N/A: not applicable.

XSL•FO
RenderX

The Accessibility, Relevance, and Impact Conceptual Framework of Digital Inclusivity for Health Information Systems

Our analysis resulted in the development of a conceptual framework of digital inclusivity for health information systems that comprises 3 essential elements: accessibility, relevance, and impact (Figure 3). Per the Accessibility, Relevance, and Impact (ARI) framework, in evaluating a health information system's accessibility, the availability and affordability or sustainability of access to the internet (for web-based applications); necessary devices (eg, computers, smartphones, and tablets); and the application or system itself must be considered. Also relevant are a system's perceivability (the extent to which it can be used by individuals with different sensory abilities, such as visual or hearing impairments), operability (the extent to which it can be used by individuals with different physical abilities or technological proficiencies), and comprehensibility (the extent to which users can understand and accurately interpret the system's content). A health information system's relevance is also key to its digital inclusivity. Digitally inclusive systems are useful (ie, they fulfill a clear purpose); trustworthy (ie, they are viewed as credible); and aligned with users' values, beliefs, customs, and preferences (ie, they are congruent). Finally, the framework suggests that the evaluation of a health information system's digital inclusivity requires consideration of its impact, that is, the extent to which it improves (or would be expected to improve) users' lives (benefit) and the presence or absence of protection from web-based threats (eg, malware and data breaches) associated with the system's use (safety).

Figure 3. The Accessibility, Relevance, and Impact (ARI) conceptual framework of digital inclusivity for health information systems.



Evaluating ENVISION's Digital Inclusivity Using the ARI Framework

Accessibility

Data describing ENVISION's availability referenced the presence or lack of internet access, technological devices, or otherwise referred to potential users' ability to retrieve the application. Study participants' feedback and experiences related to availability were generally positive. Family caregivers either did not comment on ENVISION's availability or were positive in their responses; no family caregivers identified availability-related barriers to accessing the application. Although hospice clinicians also expressed generally positive

XSL•FO
RenderX

perceptions of ENVISION's availability, some clinicians noted potential limitations. For example, one clinician stated that the application would be inaccessible for some, including "people without internet, without computers." Another explained, "Keep in mind that your patients or family members...may not have access to technology." With regard to the application's availability to clinicians, one participant emphasized the importance of integrating any new tools with the existing electronic health record:

> *Sometimes more tools are better, but sometimes more tools take more time...If this could be embedded into our current [EHR], I could find that helpful. To log out and then log in to something else or to toggle between two applications, I think, would be more cumbersome.*

Data describing ENVISION's affordability and sustainability referenced the cost of ENVISION itself or other resources (eg, internet access and technological devices) required to access and use the application. Participants who commented on ENVISION's affordability and sustainability expressed generally positive perceptions. Family caregivers, many of whom likely assumed the application would be included in routine hospice care and thus free of charge to patients and families, did not directly raise the issue. Two clinicians (one of whom also occupied an administrative role) noted the importance of making the application affordable and sustainable for hospice agencies, stressing the issue of cost-effectiveness, and linking the application to clinically relevant outcomes. When asked if they would recommend routine use of ENVISION in hospice care, they replied as follows:

> *I think you'd have to look at...cost, [but] this program will definitely improve outcomes for symptom management.*

Data describing ENVISION's perceivability referenced the extent to which the application could be used by individuals with different sensory abilities, particularly with regard to vision. Participants' comments regarding the perceivability of ENVISION varied. Some suggested that the font size was too small:

> *The print would have to be a little larger. I think, in general, anything to do with [older adults] should be larger.*

Others stated the opposite, describing ENVISION's text and images as "not too small." One participant who did not cut and paste the one-time password from their welcome email into the log-in screen noted that "for people [who] are...visually challenged, [entering] the password [could be] a bit of a headache." A clinician described ENVISION as a "wonderful option for patients and family members" but suggested that it not be required, as some may not be "able to see and hear...and all those sorts of things."

Data describing ENVISION's operability referenced the extent to which it could be used by individuals with different physical abilities or technological proficiencies. Participants provided mixed feedback on ENVISION's operability. Overall, for family caregivers, logging into the application for the first time (which required them to enter their email address and a one-time password emailed to them during the call) was more challenging than using it once they logged in. One family caregiver stated as follows:

> *If you're not a computer-savvy sort of person, [logging in the first time] could be a challenge.*

This was particularly the case for users (family caregivers and hospice clinicians) who did not cut and paste their one-time password from their welcome email into the log-in screen:

> *I have to write [the one-time password] down because I won't remember it.*

Another user provided a specific suggestion related to this issue:

> *Do you know, with this system, can you instruct people to copy the initial password and paste it in, just to make it easier for people?*

After entering the system, however, family caregivers could easily enter symptom and well-being data into ENVISION, describing this process as "pretty simple" and "really easy." One family caregiver stated, "That took less than 30 seconds," while another explained as follows:

> *I think it was really easy. I really liked how there's a definition [of each symptom or wellbeing indicator]...It's comparable to other applications I use for work...I think that would be pretty straightforward for the general population, too.*

Although this user saw the information symbol (lowercase "i" with a circle around it) located next to each symptom and well-being indicator and knew to click on it for more information, others required prompting before being able to do so. Ultimately, however, 100% of the users who expressed a desire for more information about a symptom or well-being indicator were able to successfully obtain that information by clicking on the information symbol independently or after receiving the following verbal prompt from the researcher: "Is there anything on the screen that might give you more information about that?"

Clinicians' comments regarding ENVISION's operability with regard to logging into the system and navigating the application generally mirrored those of family caregivers. They described the overall application as generally operable while emphasizing that it might be challenging for individuals to use if they were physically unable to type or "[could not] even use a smartphone" (the issue of family caregivers' ability to independently use the application is further described when discussing *congruence* under the *Relevance* section, as numerous clinicians expressed concern that they would be tasked with training and assisting technologically challenged family caregivers with ENVISION's use, requiring significant amounts of their already limited work time). Most of the unique data about ENVISION's operability for clinicians focused on using the interactive graphs that enabled longitudinal viewing of symptoms and well-being indicators (these graphs were available only to clinician users). To choose which symptoms or well-being indicators appeared on the longitudinal graphs, clinician users needed to click a box next to the appropriate symptoms or indicators, which multiple users failed to do without prompting or considerable thought.

XSL•FO
RenderX

For example, one clinician user's think-aloud data included the following:

> *I'm guessing maybe—I was looking at it on my computer—the little check boxes underneath the graph might affect the graph, I guess...Now, let's see...[begins clicking on boxes and noting changes to the graph].*

Another suggested that the application be modified to include "some education on what that graph is and how to utilize it." Other clinician users appeared to interact with the graph more intuitively and were observed easily manipulating it. One such clinician stated as follows:

> *I didn't have any problem with it. I'm middle-aged and...pretty computer-literate. I didn't have any problems with it at all.*

Data describing ENVISION's comprehensibility referenced the extent to which users could understand and accurately interpret the application's content. ENVISION's comprehensibility was determined to be mixed. Overall, family caregivers could easily comprehend ENVISION's content, successfully entering symptom and well-being indicators and accurately interpreting the data visualizations (labeled boxes shaded in different intensities to reflect symptom and indicator intensity) featured on patient and caregiver scorecards. Several described the content as easy to understand, using words and phrases such as "straightforward" and "written in plain English." Among family caregivers, comprehensibility challenges were limited to understanding the definition of specific symptoms or well-being indicators; however, most of these challenges were resolved when the users clicked on the information symbol and were shown a definition. Users commonly clicked on the information symbol next to "well-being," expressing confusion about what it entailed (eg, "Is that mental well-being or is that physical?"). Differentiating between "tiredness" and "drowsiness" was also challenging for numerous family caregivers. Among the comprehensibility challenges that were not resolved by clicking on the information symbol, nonspecificity (eg, uncertainty whether they were being asked to report on generalized anxiety or anxiety specific to the hospice experience and confusion about the insomnia indicator: "Is that insomnia [as in] you can't sleep, or is it just that you know you have to get up because you have to check [on the patient]?") was by far the most common. Clinicians recommended that longitudinal graphs be labeled with complete descriptions of symptoms and well-being indicators rather than shortened descriptors (eg, use "shortness of breath" rather than just "breath"). However, this may have been more of a design preference than an issue related to comprehensibility. One family caregiver recommended that features beyond the patient or caregiver name and uploaded photograph be included to remind the family caregiver when they were being asked to report on the patient's experience or their own:

> *You could say, "Now we're...addressing you, not the patient" or however you would want to say it...Make it clearer which page is for the patient and which is for the caregiver.*

Relevance

Data describing ENVISION's purpose referenced its perceived usefulness. Most users identified a clear and important purpose for ENVISION in their respective roles. Family caregivers repeatedly emphasized the importance of reporting symptoms and well-being indicators to the hospice team (patient pain was most commonly cited as a high-priority symptom to communicate). Feedback on the importance of the general well-being indicator, however, was mixed among family caregivers. Some family caregivers selected it as the most critical piece of data to communicate, while others saw it as redundant:

> *I feel that was a culmination of all the options that you gave me to begin with. If I'm already addressing each one of those issues individually…maybe I didn't necessarily need to rate it separately.*

One caregiver was unclear why caregiver insomnia was included as a well-being indicator:

> *If I had insomnia, how would the healthcare provider help me with that?*

With a few exceptions, clinician participants could readily identify a purpose for ENVISION in their clinical role, evident from representative statements as follows:

> *I think it would help me do my job better due to it being so precise, and I go back to the [patient and caregiver scorecard]. It's a lot easier for me to see what's going on with that patient the way that was presented than what I'm doing now in a chart, where I have to click and copy and paste and go here and there and everywhere [to] different notes and things like that.*

A chaplain explained how using ENVISION would enhance spiritual care:

> *[ENVISION might help me decide] how soon I might want to make another visit. Because if the person is very spiritual and prayer or listening to hymns or singing [helps] with pain or anxiety, [using ENVISION would allow me to] see if maybe another visit might be something that they might appreciate sooner than later.*

One chaplain user, who expressed generally positive perceptions of ENVISION's relevance, suggested that the application would be improved by the inclusion of an indicator for "some type of spiritual distress." With regard to the graph's usefulness, clinician feedback was generally positive. One clinician explained as follows:

> *[ENVISION] would be helpful to identify patterns without having to go back and read your notes, and it would also be helpful to measure how long a pattern's been happening when it might be hard to conceptualize that just through memory.*

More general comments described ENVISION as "a really cool tool and a really good idea [that would be] really useful" and "really helpful." Another stated as follows:

XSL•FO
RenderX

> *I would be eager for [ENVISION]. I think it would be great for patients' families [to feel] like they have another method of communicating with us.*

A clinician described the patient and caregiver scorecards as follows:

> *I think they are very helpful. It's easy, quick to identify, and you can see exactly what the problems are.*

A few clinicians specifically commented on including caregiver well-being indicators in addition to patient data, noting its usefulness:

> *What I really appreciate is that...it indicates an attention to continued review about how the caregiver is doing, and that isn't always done.*

A hospice physician stated as follows:

> *I think having access to this would really help, so I can get the patient perspective. As a hospice physician, a lot of times I'm getting just a third-party review from the nurse. I don't necessarily get this drilled-down of a rating scale on what's going on.*

Two clinician users were more negative than positive regarding ENVISION's purpose. One (the more ambivalent of the 2) user stated as follows:

> *I think it's helpful, but is it necessary? I don't know.*

The other user explained their reservations about the application:

> *[My] knee-jerk response to [ENVISION] is why the heck would I be looking at a computer and not talking to [the family caregiver]? I have no idea why we would add a layer between the hospice nurse and the [patient and family]...I'm struggling with the whole concept...I'm a [age in the 60-69 range]-year-old nurse, and I'm covering two different teams...[Even] with 21 patients, I would still want to have direct conversations with my patients and families. I would want them to feel like they have no barriers whatsoever to either calling the office or calling my work cell phone and saying, "Guess what's happening this morning?"*

Data describing ENVISION's congruence referenced the degree to which the application was aligned with users' values, beliefs, customs, and preferences. When discussing ENVISION's congruence, participants commented on qualities such as the application's goodness of fit or described what they liked and disliked about it. Overall, family caregivers generally reported ENVISION to be aligned with their values, beliefs, customs, and preferences. None of the family caregivers reported perceived or anticipated challenges with daily symptom and well-being data entry. Two family caregivers mentioned specific symptoms that seemed at odds with their expectations or understanding of hospice care. One questioned why the hospice team would need to know whether they (the family caregiver) were experiencing insomnia (as previously described), and the other thought asking about patients' lack of appetite might be problematic, as they understood decreased appetite to be a normal part of the dying process rather than something that required a clinical response: "I was just told, 'Don't try to make her [eat].'" Other data suggested that this family caregiver's concern might have been warranted, as one caregiver cited "lack of appetite" as among the most important symptoms to communicate to the hospice team, explaining that a hospice patient "needs to eat." While cultural congruence was infrequently discussed regarding the ENVISION application, the few comments provided were positive and related to cultural norms that might reduce the likelihood of unscheduled contacts with the hospice team in the absence of a tool such as ENVISION:

> *In an ideal world, every clinician would call [the family caregiver of a patient whose pain medications were increased] the next day to check in to see if this is working better, but I know that's not going to be the case. A lot of caregivers actually wait a full week until the...nurse comes back, and I'm like, 'Oh, don't do that. Let them know that it's working. Let them know if it's not working. Because [patients] don't need to suffer like that.' There are cultural values that limit how people communicate, and that's especially true in, like, Latino populations and other people who have been marginalized before who don't know that they're also an authority in this, in the reporting of patients' symptoms.*

Congruence pertaining to clinician data primarily related to clinicians' preferences and experiences as busy professionals with limited time to engage in additional work tasks. These data primarily addressed the provision of technical support or data entry reminders to family caregivers using ENVISION, something clinicians were almost universally disinclined to take on. In addition, the previously described response from the hospice nurse with a strong preference for nontechnologically mediated communication ("Why we would add a layer between the hospice nurse and the [patient and family]?") was identified as a likely example of perceived incongruence with the clinician's personal values (ie, an aversion to technology or belief that more traditional forms of communication are more effective or personal). Conversely, perceptions of ENVISION as a tool to increase efficiency were strongly related to perceptions of the application as a good fit for clinicians' workflow. For example, one clinician emphasized the timesaving value of ENVISION's patient and caregiver scorecards:

> *It doesn't seem like there's a lot of information on [the scorecards] that doesn't need to be there, so that's helpful...Whenever I'm reading people's [medical] records, I'm just like, "Where is the information I'm looking for?"*

Descriptions of the application as "a quick snapshot" and as allowing clinicians to quickly identify symptoms in need of attention were common.

Data describing ENVISION's credibility referenced the degree to which users perceived the application as trustworthy. Participants rarely commented on ENVISION's credibility. Furthermore, 100% of the data segments labeled with the code "credibility" were also labeled with the code "safety" and were found to pertain more directly to security issues than to the

XSL•FO
RenderX

perceived trustworthiness of the application. Thus, to avoid duplicate reporting of findings, these data are described in the context of ENVISION's safety, which is discussed in the section describing findings related to ENVISION's impact.

Impact

Data describing ENVISION's benefit referenced the ways in which the application might improve users' lives. Both family caregivers and clinicians cited potential benefits of the ENVISION application, primarily centered around better symptom management and increased awareness of opportunities to improve patients' and family caregivers' quality of life. Much of the information labeled with code "benefit" was also labeled with the code "relevance" due to perceived improvement in individuals' ability to complete tasks associated with their respective roles, whether as family caregivers or hospice clinicians. For example, a hospice clinician indicated that ENVISION "would be a good communication tool [so] that...all the team is getting the same information in real time." Another stated, "It could allow for efficient follow-up and getting the care needed to the patient probably in...a faster manner." Several clinicians predicted that ENVISION use would increase patients' and family caregivers' satisfaction with the care they received. A clinician explained as follows:

> *It would make the patients and families feel like the hospice team is more competent, that we actually work together as a team, because we would know going in [to the home] what has been going on with them for the past few days or since we've been there. I do get that a lot. Patients are like, 'I don't want to go over it again. Don't you guys talk to each other?'*

Another clinician described ENVISION as potentially empowering:

> *When your patients and families are allowed to have input, it makes them feel empowered and a part of the care. I could see how [ENVISION] would be beneficial for the patients or their families to utilize.*

One clinician user identified benefits from 3 perspectives:

> *From the caregiver's perspective, I think it's helpful to have some sense of feeling like you have an outlet to discuss what symptoms you're having so that you can actually get help from the interdisciplinary team. I think it's helpful from the patient's perspective to kind of have a sense of control over how their symptoms are being managed...I think it's helpful from the provider's perspective for…symptom management, changes in medications, gauging how they're working, as well as helping guide that family with new symptoms that are showing up and education as well as prognostication.*

Comments describing potential drawbacks of the application were less frequent and often co-coded with other digital inclusivity elements. For example, clinicians worried that family caregivers with limited technological skills might feel frustrated when interacting with the application. Clinicians also worried that family caregivers would find it burdensome to enter daily symptom assessment data:

> *[Having daily symptom and well-being data] would be sweet. That might be a big ask for some caregivers. One more thing to do.*

However, no family caregivers cited daily data entry as likely burdensome. Clinicians also cautioned against using ENVISION data to reduce or "change the care that we otherwise would attempt to provide."

Data describing ENVISION's safety referenced the presence or absence of protection from online threats associated with the application's use, such as malware or data breaches. ENVISION's safety was rarely addressed. When the users did address it, they tended to focus on password-related hassles rather than concerns that using the application made them susceptible to digital threats. One exception was that clinicians emphasized the need for any application used in hospice to be Health Insurance Portability and Accountability Act compliant, as that would likely be required for adoption into routine care.

Discussion

Principal Findings

We developed a conceptual framework of digital inclusivity for health information systems and then applied the framework in evaluating the digital inclusivity of ENVISION, a symptom monitoring tool for home hospice care. Our analysis identified accessibility, relevance, and impact as essential considerations in assessing a health system's digital inclusivity; all 3 were incorporated into our newly created ARI framework. Study findings resulting from our application of the ARI framework generally supported ENVISION's digital inclusivity, particularly concerning its perceived relevance to the work of family caregivers and hospice clinicians and its potentially positive impact on symptom management and quality of life. Limitations to ENVISION's digital inclusivity centered around issues of accessibility, particularly operability among individuals with limited technological knowledge and skills.

The ARI framework is informed by and extends prior knowledge. It incorporates constructs from several existing models, most notably the community-oriented framework on which it was explicitly based [15]. Both frameworks place a strong emphasis on accessibility, including availability, affordability, and more standard accessibility features, conceptualized in the ARI framework as perceivability, operability, and comprehensibility (these closely mirror principles detailed in the widely referenced Web Content Accessibility Guidelines 2.0, authored by the World Wide Web Consortium) [25]. The ARI framework also echoes some of the principles highlighted in usability heuristics for user interface design given by Nielsen [26] (eg, the match between the system and the real world) and Technology Acceptance Model elaborated by Davis [27] (eg, usefulness and ease of use). In building on prior research, the ARI framework synthesizes relevant constructs from numerous bodies of existing work, setting the stage for meaningful assessment of the digital inclusivity of individual tools. In addition to providing a valuable synthesis of existing models pertinent to digital inclusivity, the ARI framework incorporates constructs uniquely relevant to the context of health information systems. It

XSL•FO
RenderX

identifies patients and families, clinicians, and institutions as unique yet interdependent user types, each with potentially different cultures, responsibilities, needs, and concerns. Furthermore, it is grounded in data derived from home hospice care, a clinical context that highlights the extent to which patients and families are increasingly required to be both care providers, via family caregiving [28] and disease self-management [29], and care recipients, via patient- and family-centered models of care [30].

Importantly, the salience of specific constructs highlighted in the ARI framework will likely fluctuate over time. For example, limited internet availability may become less problematic in the United States, where the federal government's recent infrastructure investments are predicted to significantly expand rural internet availability [31]. Similarly, while health information systems' operability will likely continue to be important, tools requiring basic technological skills to operate may become more broadly operable due to demographic shifts, as the proportion of potential users who are digital natives (people who grew up regularly using digital technologies [32]) is expanding. Other issues, such as cost-related barriers to accessing digital health tools, seem likely to retain their importance over time, particularly in light of increased recognition of income inequality and other social determinants of health [33].

Limitations

Study findings should be interpreted in light of numerous limitations. First, our study sample was disproportionate in terms of having higher number of non-Hispanic, White, and female individuals. All family caregivers who participated in the study had at least some college education or trade school experience, and all could speak and read English. Furthermore, while our sample reflected some variability regarding functional abilities (eg, some participants reported mild visual impairment requiring corrective lenses), no participants reported significant physical disabilities. Additional research with more diverse participants, including individuals with varying degrees of literacy and functional ability, is needed to refine the ARI framework and better understand and ultimately enhance ENVISION's digital inclusivity. Notably, the ARI framework is in its infancy, and additional development and testing will likely be needed to maximize its potential impact on the field. In particular, noted conceptual links between relevance and the benefit subcategory of impact highlight the need for ongoing attention to issues of construct validity. With regard to ENVISION's potential for clinical adoption, recommended next steps include pilot testing in real-world scenarios, followed by more definitive efficacy testing to determine its effect on outcomes identified by users as areas of potential impact, such as symptom management and quality of life. In addition, in developing the ARI framework, we opted to include some elements of digital inclusivity even in the absence of data from the ENVISION evaluation supporting their inclusion if existing literature or expertise among team members suggested that they were essential to the concept of digital inclusivity. For example, although participants rarely discussed safety, it was retained from the original framework in light of the large and growing number of digital security threats in existence and noted disparities in individuals' knowledge of cybersecurity [34]. Thus, while the ARI framework is primarily grounded in data derived from our evaluation of ENVISION's digital inclusivity, some exceptions apply. Finally, we emphasize that all study participants used ENVISION in a hypothetical manner, interacting with data either from memory (as was the case for family caregiver participants) or from fictitious patients and caregivers (as was the case for clinician participants). We cannot conclude with certainty that individuals using the application in real-life situations would have similar experiences or provide feedback mirroring that provided by the study participants. The hypothetical nature of participants' application use also limited their ability to provide feedback on certain aspects of ENVISION, such as the application's actual costs, including the labor and other resources required to support and sustain its use. Additional research is needed to determine ENVISION's costs and its benefits to home hospice patients, family caregivers, and clinicians.

Conclusions

Our evaluation of ENVISION identified many ways by which the tool is digitally inclusive. Although specific users' experiences and feedback varied, ENVISION was determined to be generally accessible by individuals with the skills and resources required to access and operate typical web-based applications. This overall assessment was most explicitly reflected in users' comparisons of ENVISION's operability to that of applications that they regularly encountered in their work and personal lives. User data were most positive regarding ENVISION's relevance, with nearly all family caregivers and clinicians readily identifying multiple use case scenarios for the application in home hospice care. Although individuals participating in the study interacted with hypothetical patient and family caregiver data, most predicted numerous, meaningful, positive outcomes of ENVISION use, including improved symptom management and patient and caregiver well-being.

User data also provided insights into ways in which ENVISION's digital inclusivity is limited. While most Americans can access the internet from home, this capability remains limited among older adults, racial and ethnic minority groups, and individuals residing in rural communities and low-income households [35,36]. As an entirely web-based application requiring daily use, ENVISION would largely be inaccessible to individuals without high-speed internet access at home or nearby. Moreover, individuals with limited technological skills may be unable to use the application without training and support, which many hospice agencies lack the resources to provide. Minimally, our findings suggest that support would be needed to assist first-time users in logging into the system and creating a new password, as this proved to be the most challenging aspect of operating ENVISION for many users. Adoption of password alternatives (eg, biometrics, physical hardware, etc) may be considered as this technology evolves [37]. In addition, offering support in using existing tools to enhance accessibility (eg, the zoom-in feature or magnifying applications on mobile phones to enhance character visibility) may be needed. Incorporation of existing principles (eg, the usability heuristics by Nielsen) [26] into future design efforts would likely enhance operability and is thus supported by study

XSL•FO
RenderX

findings. With regard to relevance, ENVISION may be a poor fit for family caregivers and clinicians who prefer face-to-face (or telephone) contact over more technologically mediated communication. Clinicians' concerns that the application might lead to decreased face-to-face contact might be assuaged by presenting ENVISION as a tool to supplement, not substitute, in-home patient and family care. Finally, findings clearly indicate that ENVISION must provide clinicians with a net gain in terms of efficiency, consistent with existing research highlighting time constraints as the most significant professional challenge for hospice clinicians [38].

Acknowledgments

The Engagement and Visualization to Improve Symptoms in Oncology Care (ENVISION) application was created in partnership with the software engineering and user experience team at PlatformSTL. Funding for ENVISION and this study was provided by the BJC HealthCare and Washington University in St Louis Healthcare Innovation Lab and the Washington University in St. Louis Institute for Informatics, Data Science, and Biostatistics through the Big Ideas Data-Driven Innovation Competition (principal investigator: KTW) and by the Foundation for Barnes-Jewish Hospital (principal investigator: PW). All works made for hire are the exclusive joint property of Washington University in St. Louis and BJC HealthCare. The Recruitment Enhancement Core in the Network Capacity component of the Washington University Institute of Clinical and Translational Sciences is supported by the National Center for Advancing Translational Sciences of the National Institutes of Health (UL1TR002345). The paper's content is solely the authors' responsibility and does not necessarily represent the official views of the National Institutes of Health or any other funding entities.

Authors' Contributions

All authors contributed to the writing of the article and provided substantive feedback on its content. KTW designed and directed the study and led data analysis and interpretation of findings. DPO and GD contributed to data analysis and interpretation of findings. AKD served as the study coordinator, collected data, and contributed to data analysis and interpretation of findings. PGL, PW, and JJB contributed to the study design and interpretation of findings.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Original community-level digital inclusivity principles and final codes.
[DOCX File , 19 KB - humanfactors_v11i1e51789_app1.docx]

References

1. Tatum PE, Mills SS. Hospice and palliative care: an overview. Med Clin North Am 2020 May;104(3):359-373. [doi: 10.1016/j.mcna.2020.01.001] [Medline: 32312403]
2. NHPCO facts and figures. National Hospice and Palliative Care Organization. 2022 Dec. URL: https://www.nhpco.org/wp-content/uploads/NHPCO-Facts-Figures-2022.pdf [accessed 2024-05-07]
3. Casarett D, Harrold J, Harris PS, Bender L, Farrington S, Smither E, et al. Does continuous hospice care help patients remain at home? J Pain Symptom Manage 2015 Sep;50(3):297-304 [FREE Full text] [doi: 10.1016/j.jpainsymman.2015.04.007] [Medline: 25936937]
4. Haley WE, Allen RS, Reynolds S, Chen H, Burton A, Gallagher-Thompson D. Family issues in end-of-life decision making and end-of-life care. Am Behav Sci 2016 Jul 27;46(2):284-298. [doi: 10.1177/000276402236680]
5. McMillan SC. Interventions to facilitate family caregiving at the end of life. J Palliat Med 2005;8 Suppl 1:S132-S139. [doi: 10.1089/jpm.2005.8.s-132] [Medline: 16499461]
6. Washington K. Commitment to care: a grounded theory of informal hospice caregivers' development as symptom managers. Qual Soc Work 2011 Dec 29;12(3):358-371. [doi: 10.1177/1473325011429047]
7. Mather H, Kleijwegt H, Bollens-Lund E, Kelley AS, Ornstein KA. Symptom management experience of end-of-life family caregivers: a population-based study. J Pain Symptom Manage 2022 Dec;64(6):513-520 [FREE Full text] [doi: 10.1016/j.jpainsymman.2022.07.017] [Medline: 35944883]
8. Chi NC, Demiris G. Family caregivers' pain management in end-of-life care: a systematic review. Am J Hosp Palliat Care 2017 Jun;34(5):470-485. [doi: 10.1177/1049909116637359] [Medline: 26975303]
9. Ratkowski KL, Washington KT, Craig KW, Albright DL. The stress of sadness: the most stressful symptoms for hospice family caregivers. Am J Hosp Palliat Care 2015 Nov;32(7):745-749. [doi: 10.1177/1049909114540034] [Medline: 24982302]
10. Tarter R, Demiris G, Pike K, Washington K, Parker Oliver D. Pain in hospice patients with dementia: the informal caregiver experience. Am J Alzheimers Dis Other Demen 2016 Sep 09;31(6):524-529 [FREE Full text] [doi: 10.1177/1533317516653825] [Medline: 27303062]

XSL•FO
RenderX

11. Washington KT, Wilkes CM, Rakes CR, Otten SJ, Parker Oliver D, Demiris G. Relationships among symptom management burden, coping responses, and caregiver psychological distress at end of life. J Palliat Med 2018 Sep;21(9):1234-1241. [doi: 10.1089/jpm.2017.0584] [Medline: 29727256]
12. Phongtankuel V, Teresi JA, Eimicke JP, Kong JX, Adelman RD, Prigerson HG, et al. Identifying the prevalence and correlates of caregiver-reported symptoms in home hospice patients at the end of life. J Palliat Med 2020 May;23(5):635-640 [FREE Full text] [doi: 10.1089/jpm.2019.0324] [Medline: 31873053]
13. Washington KT, Demiris G, Oliver DP, Backonja U, Norfleet M, Han CJ, et al. ENVISION: a tool to improve communication in hospice interdisciplinary team meetings. J Gerontol Nurs 2020 Jul 01;46(7):9-14. [doi: 10.3928/00989134-20200605-03] [Medline: 32597996]
14. Yao R, Zhang W, Evans R, Cao G, Rui T, Shen L. Inequities in health care services caused by the adoption of digital health technologies: scoping review. J Med Internet Res 2022 Mar 21;24(3):e34144 [FREE Full text] [doi: 10.2196/34144] [Medline: 35311682]
15. Building digital communities: a framework for action. Institute of Museum and Library Services. URL: https://www.imls.gov/publications/building-digital-communities-framework-action [accessed 2024-05-07]
16. van Kessel R, Roman-Urrestarazu A, Anderson M, Kyriopoulos I, Field S, Monti G, et al. Mapping factors that affect the uptake of digital therapeutics within health systems: scoping review. J Med Internet Res 2023 Jul 25;25:e48000 [FREE Full text] [doi: 10.2196/48000] [Medline: 37490322]
17. Sieck CJ, Sheon A, Ancker JS, Castek J, Callahan B, Siefer A. Digital inclusion as a social determinant of health. NPJ Digit Med 2021 Mar 17;4(1):52 [FREE Full text] [doi: 10.1038/s41746-021-00413-8] [Medline: 33731887]
18. Portz JD, Fruhauf C, Bull S, Boxer RS, Bekelman DB, Casillas A, et al. "Call a teenager… that's what I do!" - Grandchildren help older adults use new technologies: qualitative study. JMIR Aging 2019 Jun 06;2(1):e13713 [FREE Full text] [doi: 10.2196/13713] [Medline: 31518274]
19. Usability. U.S. General Services Administration. URL: https://www.usability.gov/how-to-and-tools/resources/ucd-map.html [accessed 2017-01-08]
20. Mik-Meyer N. Multimethod qualitative research. In: Silverman D, editor. Qualitative Research. London, UK: SAGE Publications; 2020.
21. Santhosh L, Rojas JC, Lyons PG. Zooming into focus groups: strategies for qualitative research in the era of social distancing. ATS Sch 2021 Jun;2(2):176-184 [FREE Full text] [doi: 10.34197/ats-scholar.2020-0127PS] [Medline: 34409412]
22. Lundgrén-Laine H, Salanterä S. Think-aloud technique and protocol analysis in clinical decision-making research. Qual Health Res 2010 Apr;20(4):565-575. [doi: 10.1177/1049732309354278] [Medline: 19959822]
23. Lunde P, Skoglund G, Olsen CF, Hilde G, Bong WK, Nilsson BB. Think aloud testing of a smartphone app for lifestyle change among persons at risk of type 2 diabetes: usability study. JMIR Hum Factors 2023 Nov 15;10:e48950 [FREE Full text] [doi: 10.2196/48950] [Medline: 37966894]
24. Fisher G, Aguinis H. Using theory elaboration to make theoretical advancements. Org Res Method 2017 Feb 03;20(3):438-464. [doi: 10.1177/1094428116689707]
25. Web Content Accessibility Guidelines (WCAG) 2.0. World Wide Web Consortium. 2008. URL: https://www.w3.org/TR/WCAG20/ [accessed 2023-08-11]
26. Nielsen J. Enhancing the explanatory power of usability heuristics. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. 1994 Presented at: CHI '94; April 24-28, 1994; Boston, MA. [doi: 10.1145/191666.191729]
27. Davis FD. Perceived usefulness, perceived ease of use, and user acceptance of information technology. MIS Q 1989 Sep;13(3):319-340. [doi: 10.2307/249008]
28. Washington KT. Hospice and caregiving. In: Gu D, Dupre ME, editors. Encyclopedia of Gerontology and Population Aging. Cham, Switzerland: Springer; 2021.
29. Bodenheimer T, Lorig K, Holman H, Grumbach K. Patient self-management of chronic disease in primary care. JAMA 2002 Nov 20;288(19):2469-2475. [doi: 10.1001/jama.288.19.2469] [Medline: 12435261]
30. Rawson JV, Moretz J. Patient- and family-centered care: a primer. J Am Coll Radiol 2016 Dec;13(12 Pt B):1544-1549. [doi: 10.1016/j.jacr.2016.09.003] [Medline: 27888939]
31. Biden-Harris administration announces over $700 million to connect people in rural areas to high-speed internet. U.S. Department of Agriculture. 2023 Jun 12. URL: https://www.usda.gov/media/press-releases/2023/06/12/biden-harris-administration-announces-over-700-million-connect [accessed 2023-07-05]
32. Prensky M. Digital natives, digital immigrants part 1. On the Horizon 2001 Sep;9(5):1-6. [doi: 10.1108/10748120110424816]
33. Truesdale BC, Jencks C. The health effects of income inequality: averages and disparities. Annu Rev Public Health 2016;37:413-430. [doi: 10.1146/annurev-publhealth-032315-021606] [Medline: 26735427]
34. Vogels EA, Anderson M. Americans and digital knowledge in 2019. Pew Research Center. 2019. URL: https://www.pewresearch.org/internet/2019/10/09/americans-and-digital-knowledge/ [accessed 2023-06-01]
35. Internet/broadband fact sheet. Pew Research Center. 2024 Jan 31. URL: https://www.pewresearch.org/internet/fact-sheet/internet-broadband/ [accessed 2024-05-13]

XSL•FO
RenderX

36. Swenson K, Ghertner R. People in low-income households have less access to internet services. Office of the Assistant Secretary for Planning & Evaluation, U.S. Department of Health & Human Services. 2020 Apr. URL: https://aspe.hhs.gov/sites/default/files/private/pdf/263601/Internet_Access_Among_Low_Income.pdf [accessed 2023-07-14]
37. van Oorschot PC. User authentication—passwords, biometrics and alternatives. In: Computer Security and the Internet. Information Security and Cryptography. Cham, Switzerland: Springer; 2021.
38. Washington KT, Alaniz Staner LM, Collie BE, Craig KW, Demiris G, Oliver DP. Toward a practice-informed agenda for hospice intervention research: what are staff members' biggest challenges? Am J Hosp Palliat Care 2021 May;38(5):467-471. [doi: 10.1177/1049909120948222] [Medline: 32757823]

Abbreviations

ARI: Accessibility, Relevance, and Impact
ENVISION: Engagement and Visualization to Improve Symptoms in Oncology Care

Edited by A Kushniruk; submitted 11.08.23; peer-reviewed by D Chrimes, V Provencher; comments to author 18.01.24; revised version received 13.02.24; accepted 07.04.24; published 23.05.24.

Please cite as:
Washington KT, Oliver DP, Donehower AK, White P, Benson JJ, Lyons PG, Demiris G
Accessibility, Relevance, and Impact of a Symptom Monitoring Tool for Home Hospice Care: Theory Elaboration and Qualitative Assessment
JMIR Hum Factors 2024;11:e51789
URL: https://humanfactors.jmir.org/2024/1/e51789
doi:10.2196/51789
PMID:38781581

©Karla T Washington, Debra Parker Oliver, Allison K Donehower, Patrick White, Jacquelyn J Benson, Patrick G Lyons, George Demiris. Originally published in JMIR Human Factors (https://humanfactors.jmir.org), 23.05.2024. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on https://humanfactors.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Usability, Acceptability, and Preliminary Effectiveness of a Peer-Delivered and Technology-Supported Mental Health Intervention for Family Caregivers of People With Dementia: Field Usability Study

Caroline Collins-Pisano[1,2], BA; Amanda N Leggett[3], PhD; David Gambee[4], BA; Karen L Fortuna[4], PhD, LICSW

[1]Department of Psychological and Brain Sciences, Dartmouth College, Hanover, NH, United States
[2]Department of Psychology, University of Colorado, Colorado Springs, Colorado Springs, CO, United States
[3]Institute of Gerontology, Wayne State University, Detroit, MI, United States
[4]Department of Psychiatry, Dartmouth College, Hanover, NH, United States

Corresponding Author:
Caroline Collins-Pisano, BA
Department of Psychology
University of Colorado, Colorado Springs
1420 Austin Bluffs Pkwy
Colorado Springs, CO, 80918
United States
Phone: 1 (719) 255 8227
Email: ccolli20@uccs.edu

Abstract

Background: Family caregivers of people with dementia are critical to the quality of life of care recipients and the sustainability of health care systems but face an increased risk of emotional distress and negative physical and mental health outcomes.

Objective: The purpose of this study was to examine the usability, acceptability, and preliminary effectiveness of a technology-based and caregiver-delivered peer support program, the Caregiver Remote Education and Support (CARES) smartphone or tablet app.

Methods: A total of 9 adult family caregivers of people with dementia received the CARES intervention, and 3 former family caregivers of people with dementia were trained to deliver it. Quantitative data were collected at baseline and at the end of the 2-week field usability study. Qualitative data were also collected at the end of the 2-week field usability study.

Results: The field usability study demonstrated that a 2-week peer-delivered and technology-supported mental health intervention designed to improve burden, stress, and strain levels was experienced by former and current family caregivers of people with dementia as acceptable. Current family caregivers rated CARES as above average in usability, whereas the caregiver peer supporters rated CARES as marginally usable. CARES was associated with non–statistically significant improvements in burden, stress, and strain levels.

Conclusions: This field usability study demonstrated that it is possible to train former family caregivers of people with dementia to use technology to deliver a mental health intervention to current family caregivers of people with dementia. Future studies would benefit from a longer trial; a larger sample size; a randomized controlled design; and a control of covariables such as stages of dementia, years providing care, and severity of dementia symptoms.

(JMIR Hum Factors 2024;11:e41202) doi:10.2196/41202

KEYWORDS

family caregivers; dementia; peer support; technology; mobile phone

XSL•FO
RenderX

Introduction

Background

Family caregivers of people with dementia are critical to the quality of life of care recipients and to the sustainability of health care systems. Family caregivers provide US $257 billion in unpaid care to people living with dementia but face an increased risk of emotional distress and negative physical (eg, heart disease and hypertension) and mental (eg, depression and anxiety) health outcomes [1,2]. While positive gains are reported in the caregiving role, caregivers are more likely than their noncaregiving peers to report stress, burden, and strain. Approximately 46% of family caregivers of people with dementia are classified as having high levels of burden [1,2]. Burden is defined as the psychological, physical, emotional, and social challenges that family caregivers experience in response to the demands of providing care [3]. Caregivers with high levels of burden report more physical and psychological symptoms, use prescription medications and health care more frequently, and provide poorer quality of care to recipients, leading to an increased likelihood of premature institutional care [4,5]. In addition, 59% of family caregivers of people with dementia rate the emotional stress of caregiving as high or very high, and 38% rate the physical stress of caregiving as high or very high [1,2]. Stress is defined as individuals' emotional or physical responses to the challenges of caregiving, such as fatigue [6,7]. The stress of providing care to a family member with dementia increases caregivers' risk of health complications, increases their susceptibility to diseases such as hypertension, and negatively affects their quality of sleep [1,2]. Family caregivers have also reported greater levels of strain compared to caregivers of people without dementia. Strain is defined as caregivers' perception of the challenges of caregiving and their state of well-being [7]. Family caregivers who perceived themselves as having higher strain levels due to caregiving responsibilities were at a higher risk of death than those who perceived little or no strain [1,2].

Technology-Based Interventions for Caregivers

While psychosocial interventions have been shown in fully powered randomized controlled trials to reduce caregiver burden and delay nursing home admission for the care recipient, a recent meta-analysis by Walter and Pinquart [8] found that current interventions had a disappointingly small to moderate effect on caregiver well-being, burden, depression, and anxiety [4,8-10]. In addition, uptake of these interventions in the real world is limited due to caregiving obligations (between 69 and 117 hours of informal care are provided to people with dementia per week), geographical distance from the intervention, requirements to meet in person, and failure to address the personalized needs of the family caregivers.

Technology-based interventions may offset these challenges through improved accessibility of psychosocial interventions at any time or location [11,12]. A scoping review on existing technology-based peer support interventions for family caregivers found that web-based programs include websites that offer educational materials with the option to contact other informal caregivers, web-based portals with psychoeducation and peer-to-peer contact, asynchronous e-learning platforms, internet support forums and chat rooms, videoconferencing support groups facilitated by a health professional, and live virtual reality support groups facilitated by psychologists [11]. Online informal peer support groups for informal caregivers have shown high levels of engagement. Technology-based interventions enable caregivers to participate even if they are unable to leave the person with dementia unsupervised. For example, during the COVID-19 pandemic, telephone-, videoconference-, and chat room–based online support groups were the only media accessible to caregivers, a benefit due to the 24/7 nature of supporting a family member with dementia and the consequential challenges in accessing support [11,13]. Online informal peer support groups are an effective method of asynchronous web-based delivery when offered in combination with a structured psychosocial and educational intervention by skilled clinical professionals [11]. However, while some studies such as that by Han et al [14] have shown significant reductions in depression, stress, and helplessness, others such as the study by Marziali and Garcia [15] have shown only moderate effects in reducing burden and depression and increasing caregiver knowledge.

Significance of Technology-Delivered Caregiver Peer Support

To date, technology-based interventions have relied on a skilled workforce of geriatric mental health professionals. However, there are insufficient numbers of adequately trained geriatric mental health care providers [16]. As such, task shifting services from skilled clinical professionals to community members provides an emerging workforce of peer support workers (ie, interventionists without formal mental health education but with life experiences similar to those of the people they serve) [17]. Although the need for traditional clinical professionals remains, peer support services for individuals with mental health conditions have been shown to increase service accessibility without impacting service quality [17]. However, there is limited knowledge of caregiver-delivered peer support.

The use of caregiver-delivered digital peer support may promote the uptake of psychosocial interventions, reduce burden, and delay nursing home admission. Social and behavioral theories such as social support, experiential knowledge, and the helper therapy principle highlight how peer supporters have the unique ability to offer acceptance, understanding, and validation to the individuals they work with [18]. Because of their shared lived experiences, peers are often viewed as more credible and trustworthy than other health care providers and, therefore, encourage increased digital health engagement. Individuals are motivated to achieve their mental health goals (eg, reductions in burden, strain, and stress) because of the reciprocal accountability offered and modeled by their peers. Former family caregivers of people with dementia have the knowledge and skillset to deliver trained peer support to current family caregivers and could potentially benefit emotionally from the delivery of support. While there are positive outcomes associated with the end of caregiving, when family caregivers of people with dementia become former family caregivers, the detrimental effects of previous caregiving fail to improve [19]. Many former dementia caregivers experience persistent sleep disturbances,

XSL•FO
RenderX

depression, anxiety, increased physician office visits, declining health, and feelings of guilt and regret [20]. Despite evidence that former caregivers who pursue new caring roles benefit emotionally, payers and health care providers have not hired and trained former caregivers to provide evidence-based digital interventions [20].

Research Aims

The purpose of this study was to examine the usability, acceptability, and preliminary effectiveness of a technology-based and caregiver-delivered peer support program, the Caregiver Remote Education and Support (CARES) smartphone or tablet app. In this study, 3 former family caregivers of people with dementia (caregiver peer supporters) were trained in the delivery of caregiver peer support and delivered the CARES intervention and app to 9 current family caregivers of people with dementia (caregiver participants) in a 2-week field usability study.

Design of the CARES App

The CARES app and intervention were adapted from the PeerTECH smartphone or tablet app and developed to facilitate the examination of the usability, acceptability, and preliminary effectiveness of the first technology-based and caregiver-delivered peer support program in a 2-week field usability study. The PeerTECH system was designed using universal design principles and for lay interventionists (not skilled) to deliver fidelity-adherent evidence-based interventions. The PeerTECH system has been successfully used with Certified Peer Specialists, home health aides, and Certified Older Adult Peer Specialists [21]. Certified Peer Specialists are people with a mental health diagnosis who are hired, trained, and certified to provide peer support services to individuals with a similar diagnosis [22].

PeerTECH was built on the stress-vulnerability model. According to the stress-vulnerability model, the outcomes of a mental health condition are connected to biological vulnerability, stress, and protective factors (eg, peer support) [23]. PeerTECH was designed to empower individuals to address the stress and vulnerability that lead to worsening medical, psychiatric, and social health conditions. Peer specialists are trained to deliver evidence-based practices through the PeerTECH app to help participants decrease stressors and increase protective factors.

The PeerTECH app was codeveloped with peer specialists and includes 2 features developed through community-engaged research. The first is a peer-facing smartphone or tablet app that guides peers in delivering evidence-based health self-management skill development interventions. The app includes prompts to share their lived experiences of health challenges and solutions (ie, peer support) and structured intervention delivery through scripted, evidence-based training on topics such as coping skills, psychoeducation, medical management, social skills, and self-advocacy. The second feature is a participant-facing app that offers self-management support through features such as a library of self-management resources (eg, peer-led videos) and a secure messaging feature to connect with their assigned peer specialists and reinforce goals [21].

Similar to PeerTECH, the CARES app includes 2 features: a peer-facing smartphone or tablet app and a participant-facing smartphone or tablet app (Multimedia Appendix 1 for illustrations of the CARES app). The content of the CARES app, similar to that of PeerTECH, was designed according to the stress-vulnerability model to help participants decrease stressors and increase protective factors. The CARES app and intervention were adapted from the PeerTECH app by a team of researchers with expertise in peer support to guide caregiver peers in the delivery of evidence-based mental health interventions and designed according to the techniques and principles of peer support and acceptance and commitment therapy (ACT). The participant-facing CARES app offers a library of resources designed to educate participants on psychological skills (eg, mindfulness) that empower them to manage difficult thoughts (eg, acceptance) and engage in activities and behaviors that are guided by their life values and boost their well-being (eg, setting goals and identifying values) [24]. Caregiver peer supporters were trained and educated on topics such as values, goal setting, acceptance, avoidance, and negative thoughts and connected with their assigned participants via a secure messaging feature to reinforce ACT principles, share their lived experience of caregiving challenges, and share practices to enhance caregivers' wellness and mental health in relation to caregiver-related stressors [25]. The mutual practice of goal setting, for example, can help link caregivers' values to concrete plans for behavior change [26]. Figure 1 illustrates the CARES app.

The participant-facing app includes access to direct messaging with an assigned caregiver peer supporter, goals, wellness, surveys, and a resource library (see the bottom right of Figure 1). The caregiver peer supporter app allows caregiver peer supporters to message assigned participants directly (see the bottom left of Figure 1), view participants' goals and wellness plans, and view their progress in the library resource feature. The principal investigator (PI) had access to data on the participants' and peer supporters' rate of engagement with library resources, messaging, and goals and wellness features (see the top of Figure 1).

XSL•FO
RenderX

Figure 1. Caregiver Remote Education and Support app.

Description of the CARES App

Fundamentally, CARES is a peer-instructed and mediated caregiver support program that uses a smartphone app–based mechanism for communication. The CARES app consists primarily of two features: (1) a former family caregiver (caregiver peer supporter)–facing app on a smartphone or tablet that includes the option to message or video chat with the current family caregivers (participants) they have been assigned to provide caregiver peer support (Figure 2) and (2) a participant-facing app that offers the option to review an on-demand library of educational resources and a HIPAA (Health Insurance Portability and Accountability Act)-compliant text and video chat feature to communicate with their assigned caregiver peer supporter (Figure 3). The participants and caregiver peer supporters also have the option to set goals and create wellness plans. Figure 4 shows the features seen by both the participants and caregiver peer supporters.

Figure 2A depicts the home page of the caregiver peer supporter–facing CARES app. The home page includes the option to select the individual's availability to offer caregiver peer support to their assigned participants (Figure 2B), set goals, and access information on their assigned participants and chats. Figure 2C shows the options to message and video chat with the assigned participant and track their goals and wellness plan.

Figure 3 depicts the home page of the participant-facing CARES app. The home page includes the option to send messages to the assigned caregiver peer supporter, video chat directly with the assigned caregiver peer supporter, set goals, create a wellness plan, and access a library of resources (Figure 5).

XSL•FO
RenderX

Figure 2. Screenshot A depicts the home page on the caregiver peer supporter facing CARES application. The homepage includes the option to select the individual's availability to offer caregiver peer support to their assigned participants (see Screenshot B), set goals, and access information on their assigned participants and chats. Screenshot C provides options to message and video chat the assigned participant and track their goals and wellness plan.

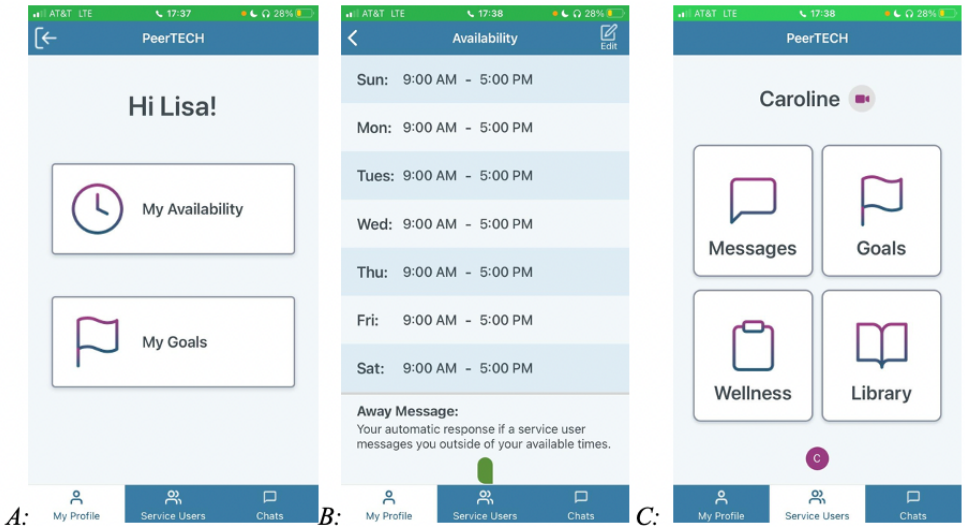
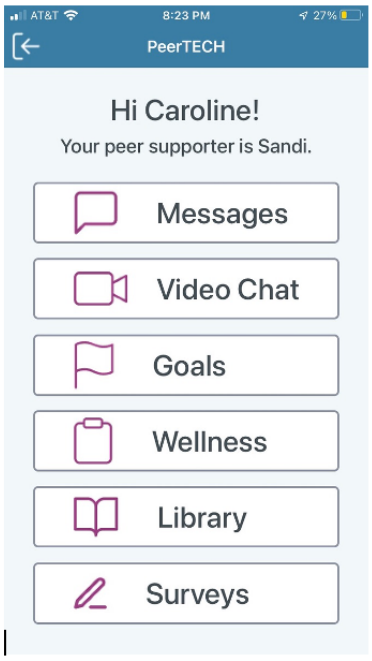


Figure 3. Caregiver Remote Education and Support participant-facing app.



The left panel in Figure 4 depicts the messaging option provided within the CARES app. Within the messaging section of the app, participants and their assigned caregiver peer supporters can send each other text-like messages. The panel in the center depicts the wellness plan. Within the wellness plan section, participants and caregiver peer supporters can add activities and strategies they wish to complete to enhance their wellness. The panel on the right depicts the goals section. Within the goals section, participants and caregiver peer supporters can set goals for themselves.

XSL•FO
RenderX

Figure 4. Caregiver Remote Education and Support features on both the participant- and caregiver peer supporter–facing apps.

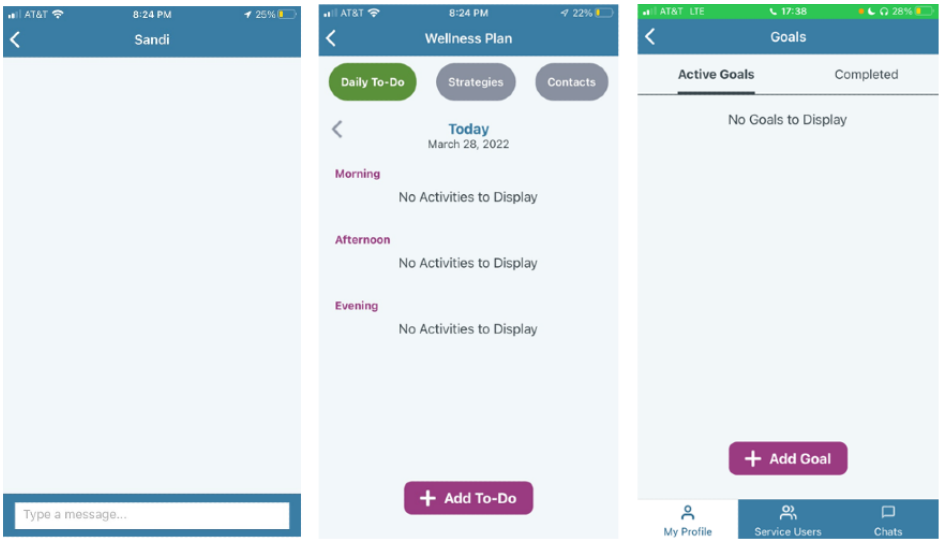
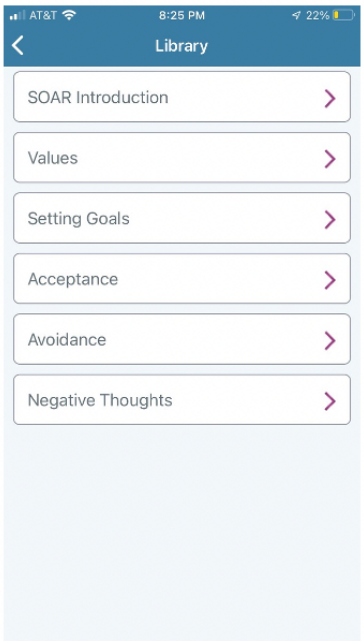


Figure 5. Caregiver Remote Education and Support library of resources.



Description of the CARES Resource Library

The CARES app includes a resource library with materials related to the principles of peer support and evidence-based practices and skills, such as ACT, to manage stress and promote mental health and well-being. CARES educational resources are designed to be reviewed by a participant and caregiver peer supporter together or individually (Figure 5). Each resource includes a combination of videos and text.

Figure 5 depicts the library of resources found in the CARES app. Each topic includes a scripted curriculum with evidence-based practices to improve mental health and well-being.

Methods

Overview

A field usability study was conducted in April 2022 to evaluate the usability, acceptability, and preliminary effectiveness of CARES as an assistive tool for guiding former family caregivers of people with dementia (n=3) in fidelity-adherent delivery to current family caregivers of people with dementia (n=12). The purpose of a field usability study is to assess the feasibility and acceptability of technology in users' natural and conceptual environments [27]. Through field usability studies, researchers gain an understanding of the problems users encounter while using the system and gain insights into how individuals use the product [27]. The field usability study was conducted for 2 weeks to provide caregiver peer supporters and participants with the time to use and assess the CARES app and establish a peer connection. The field usability study was a first step in

XSL•FO
RenderX

assessing the usability and acceptability of CARES, the study design, and the training. Caregiver peer supporters provided 5 to 7 hours of peer support per week, including video chats, messaging, and supervision. Each caregiver peer supporter was assigned 4 participants. In total, 17% (2/12) of the participants dropped out of the study due to a delay in the start date, and 8% (1/12) of the participants were excluded for not using the app and failing to initiate the field usability testing process, resulting in a final sample of 9 current family caregivers and 3 caregiver peer supporters. Study measures of burden, stress, and strain levels were administered via Qualtrics (Qualtrics International Inc) at baseline and at the end of the 2-week field usability study. Study measures on usability and acceptability were administered in an hour-long HIPAA-compliant videoconference semistructured interview at the end of the 2-week field usability study. All assessments were conducted by the PI.

Ethical Considerations

The Committee for the Protection of Human Subjects at the Dartmouth Health Institutional Review Board approved the project (FP00002112). Participants provided their written informed consent. Participants were compensated for taking part in the study. Current family caregivers received US $20 for the baseline assessment, US $20 for the completion of the 2-week CARES field usability study, and US $20 for the completion of the semistructured interview conducted after the 2-week field usability study. Caregiver peer supporters received US $120 for the 6 hours of training, US $30 per hour for the 2-week field usability study, and US $30 for the completion of the semistructured interview conducted after the 2-week field usability study.

Participants

A total of 12 participants were recruited from the Dartmouth-Hitchcock Aging Resource Center, memory cafés, and senior centers across New England and via dementia caregiver Facebook support groups to participate in the study with the goal of recruiting between 10 and 20 participants. Previous research has found that 10 users report approximately 80% of usability problems and 20 users report approximately 95% of usability problems for a given product [28]. The pilot study included 9 current family caregivers of people with dementia. Participants were eligible if they (1) were a current family caregiver of an individual with dementia, (2) were aged ≥18 years, (3) spoke and read English, and (4) provided voluntary informed consent for participation in the study. The study also included 3 former family caregivers of people with dementia. Participants were eligible if they (1) were a former family caregiver of an individual with dementia, (2) were aged ≥18 years, (3) spoke and read English, (4) were willing to use technology to deliver an intervention, and (5) provided voluntary informed consent for participation in the study. All participants were excluded if they (1) had a chart diagnosis of dementia or documented cognitive impairment as indicated by a Mini-Mental State Examination score of <24; (2) had major visual, hearing, or motor impairment; (3) had a terminal illness expected to result in death within 1 year; or (4) were patients with ≥2 emergency room visits or hospitalizations in the previous 6 months or determined by the clinical team to be psychiatrically or medically unstable.

Measures

The usability of the CARES app was evaluated using the System Usability Scale (SUS). The SUS is a widely used, valid, reliable 10-item scale that assesses system usability [29]. Sample questions include “I think that I would like to use this system frequently” and “I thought the system was easy to use.” Response options range from 1 (*strongly disagree*) to 5 (*strongly agree*). Scores range from 0 to 100, with higher scores indicating better usability [30]. A mean SUS of ≥68 is indicative of an above-average user experience [30]. A system with an SUS score of >70 is considered acceptable. Scores of >85 are considered “excellent,” scores between 71 and 84 are considered “good,” and scores between 51 and 70 are considered “OK” [30].

Caregiver burden was assessed using the Zarit Burden Interview–Short Form. The Zarit Burden Interview–Short Form is a 12-item scale that evaluates caregivers’ physical burden, financial burden, interpersonal burden, and health [3]. The Zarit Burden Interview–Short Form is a valid scale for evaluating burden in caregivers of community-dwelling individuals with dementia [3]. Sample questions include “do you feel that because of the time you spend with your relative that you don’t have enough time for yourself” and “do you feel that you have lost control of your life since your relative’s illness.” Response options range from 0 (*never*) to 4 (*nearly always*). Scores range from 0 to 48, with higher scores indicating higher levels of burden.

The Modified Caregiver Strain Index is a 13-item scale that was used to assess strain and its consequences on caregivers’ overall health. The Modified Caregiver Strain Index is a stable and reliable measure of strain among long-term caregivers [7]. Sample questions include “caregiving is inconvenient” and “I feel completely overwhelmed.” Response options include “yes, on a regular basis,” “yes, sometimes,” and “no.” Scores range from 0 to 26, with higher scores indicating higher levels of strain.

Caregiver stress was assessed using the Caregiver Self-Assessment Questionnaire (CSAQ). The CSAQ is an 18-item scale that assesses the stress levels of family caregivers [31]. The CSAQ is a valid scale for individuals caring for people with dementia [31]. Sample questions include “during the past week or so, I have felt that I couldn’t leave my relative alone” and “during the past week or so, I have been satisfied with the support my family has given me.” Response options for questions 1 to 16 include “yes” or “no.” Caregivers are considered to have high levels of stress if they respond with “yes” to ≥10 questions. Question 17 asks the caregiver to rate their level of stress from “not stressful” to “extremely stressful” on a scale from 1 to 10. Caregivers are considered to have high levels of stress if they score ≥6 on question 17.

A semistructured interview was administered to assess the acceptability of the CARES app from the perspective of the participants. The interview questions were informed by the Consolidated Framework for Implementation Research (CFIR).

XSL•FO
RenderX

The CFIR is a meta-theoretical framework that guides implementation research and is used to systematically assess potential barriers to and facilitators of implementing an intervention [32,33]. Previous studies on the feasibility of web-based tools and interventions have used the CFIR to guide qualitative analysis, identify aspects of implementation feasibility, and determine areas of improvement and adaptation to better meet the needs of users (eg, see the study by Lawson et al [34]). The PI used the CFIR Interview Guide Tool to develop the interview questions. Interview questions covered CFIR domains such as intervention characteristics (what key attributes of the intervention influence the success of implementation), patients' needs and resources (the extent to which patient needs and barriers to and facilitators of meeting those needs are accurately known), implementation climate (shared receptivity of involved individuals to an intervention and the extent to which the use of that intervention will be supported), self-efficacy (individuals' beliefs in their own capacity to successfully implement the intervention), and evidence strength and quality (individuals' perceptions of the quality and validity of the intervention) [32]. Interview guide questions included the following: "what would you change about the CARES system and intervention?" "How well does the intervention align with your values and norms?" "What barriers will family caregivers of people with dementia face to delivering or participating in the intervention?"

Procedures

Recruitment

The PI (CCP) met with staff members at the Aging Resource Center to discuss the purpose of the study and the recruitment process. Together, they identified potential groups within and outside the Dartmouth-Hitchcock Aging Resource Center to recruit both former family caregivers of people with dementia to be trained in the delivery of the CARES app and current family caregivers of people with dementia to receive the CARES intervention. If the potential participants met the inclusion criteria, they were contacted via email and provided with a description of the study. Participants were told that the study was for an honors psychology thesis that aimed to assess the usability, acceptability, and preliminary effectiveness of a technology-based and caregiver-delivered peer support program, CARES, intended to help current family caregivers of people with dementia manage burden, strain, and stress levels. If they were interested in the study, they agreed to meet with the PI via HIPAA-compliant videoconferencing software or telephone for a baseline interview. The baseline interview lasted approximately 20 to 60 minutes. For the baseline interview, the PI read through the informed consent forms and answered participants' questions regarding the content of the study. After the baseline interview, participants who provided informed consent for the study were sent a copy of their informed consent form; sent a link to a baseline survey on Qualtrics with questions on demographic information and their current burden, stress, and strain levels; and provided with information on how to download and log in to the CARES app.

Training

Once 3 former caregivers of people with dementia were recruited, they completed 6 hours of CARES training over 2 days through HIPAA-compliant videoconferencing software. The CARES training was adapted from the Digital Peer Support Certification [35]. Fortuna et al [35,36] found that a combination of educational training (the Digital Peer Support Certification) and management of the PeerTECH system increased peer support specialists' capacity to use features of the digital peer support technology [35]. The training was based on adult learning theory and experiential learning theory. Adult learning theory suggests that adults learn best when they use past lived experiences and past developed skills and knowledge to enhance their learning process [37]. Experiential learning theory consists of four principles: (1) concrete experience, (2) observation and reflection, (3) abstract conceptualization, and (4) active experimentation [38]. In the CARES training, former caregivers were taught new skills; asked to reflect on and connect new knowledge and skills to past experiences and situations; and, finally, asked to practice the new skills they had learned. Techniques such as reinforcement, summation, and teach-back were used in the CARES training to promote the mastery of peer support skills [35].

The CARES training focused on instructing the former caregivers on the basic principles and competencies in the delivery of digital peer support and evidence-based practices to manage stress. The training included an overview of the following topics: peer support, health and aging, engaging older service users with technologies, teaching older adults how to use technology, life review, acceptance, mindfulness, coaching and making a plan of action, recognizing negative thoughts, the art and science of adult learning theory, and the role of family and caregivers in technology. Facilitated group discussions were paired with a PowerPoint (Microsoft Corp) presentation. The PowerPoint presentation was provided to all caregiver peer supporters at the end of the training. After the 6-hour training, participants who provided informed consent for the study were sent a copy of their informed consent form, a link to a baseline survey on Qualtrics with questions on demographic information, a copy of the caregiver peer support training, and information on how to download and log in to the CARES app. The PI was available for technological support from Monday to Saturday between the hours of 9 AM and 5 PM.

Fidelity

Over the course of the 2-week field usability study, a member of the research team tracked the CARES app messages between the caregiver peer supporters and assigned participants to evaluate whether the caregiver peer supporters were providing peer support in line with the training.

Informed Consent

Before the 2-week field usability study, the participant was provided with a description of the study, shown the CARES app, and read aloud the consent form word for word. Participants were evaluated according to the study criteria. If the participant was eligible and provided informed consent to take part in the study, they then completed the baseline survey on Qualtrics.

XSL•FO
RenderX

Quantitative Analyses

Descriptive statistics were used to describe the demographic characteristics of the study sample and the results of the SUS. A paired-sample 2-tailed *t* test was conducted to assess the difference between the baseline and 2-week burden, stress, and strain level scores for statistical significance. All incomplete survey responses were excluded from the analyses. Descriptive statistics and analyses were computed using the SPSS software (version 26; IBM Corp).

Qualitative Analyses

Interview data were analyzed using the rigorous and accelerated data reduction (RADar) technique. The RADar method helps streamline the process of qualitative data analysis through its ability to organize, reduce, and analyze data in user-friendly software packages such as Excel (Microsoft Corp) [39]. In accordance with the RADar methodology, the interview transcripts were formatted into an all-inclusive Excel spreadsheet. The Excel spreadsheet column headings included participant number, question, response, code, and notes. One researcher assigned codes to each response. A priori codes and themes related to the CFIR framework were identified. These codes included the acceptability of CARES, user needs and resources, intervention characteristics key to the success of the implementation, self-efficacy, quality and validity of the intervention, and receptivity of users. Codes were derived from the interview data by carefully reviewing the transcribed text. The all-inclusive data table was then reduced to include only content that answered the overarching research questions and was of primary interest to the analysis. The remaining text and codes were then organized into themes that applied the CFIR framework and were adjusted to best fit the content covered in the qualitative interviews. The percentage of each theme was determined by dividing the frequency with which the theme was present in the interview quotes by the total number of interview quotes.

Results

Overview

Table 1 presents the sociodemographic characteristics of the sample at baseline. The sample of current family caregivers (9/12, 75%; the participants) had a mean age of 67.3 (SD 15.1) years and was predominantly female (6/9, 67%) and White (9/9, 100%), and most of them (4/9, 44%) cared for a spouse. The sample of former family caregivers (3/12, 25%; the caregiver peer supporters) had a mean age of 68.3 (SD 11.0) years and was predominantly female (3/3, 100%) and White (3/3, 100%). One of the caregiver peer supporters had experience caring for a parent with dementia (1/3, 33%), one had experience caring for a spouse with dementia (1/3, 33%), and the other had experience caring for a sibling with dementia (1/3, 33%).

Table 1. Sociodemographic characteristics of study participants (N=12).

Characteristics	Participants (n=9)	Caregiver peer supporters (n=3)
Sex, n (%)		
Male	3 (33)	0 (0)
Female	6 (67)	3 (100)
Age (years), mean (SD; range)	67.3 (15.1; 42-87)	68.3 (11.0; 61-81)
Race (White), n (%)	9 (100)	3 (100)
State, n (%)		
Connecticut	1 (11)	1 (33)
Florida	1 (11)	0 (0)
Massachusetts	2 (22)	1 (33)
New Hampshire	3 (33)	0 (0)
Vermont	2 (22)	1 (33)
Smartphone owner, n (%)		
Yes	9 (100)	3 (100)
No	0 (0)	0 (0)
Relation to relative with dementia, n (%)		
Child	3 (33)	1 (33)
Parent	1 (11)	0 (0)
Sibling	1 (11)	1 (33)
Spouse	4 (44)	1 (33)

A total of 3 participants were excluded from the study. In total, 33% (1/3) of these participants did not complete the 2-week CARES app field usability study and interview session. A total of 67% (2/3) of these participants did not complete the 2-week CARES app due to a delay in the 2-week field usability study with one of the caregiver peer supporters. The caregiver peer

XSL•FO
RenderX

supporter delayed the start of their 2-week field usability study because the CARES app was not functioning on their smartphone. The remaining 9 participants and all 3 caregiver peer supporters completed the CARES intervention.

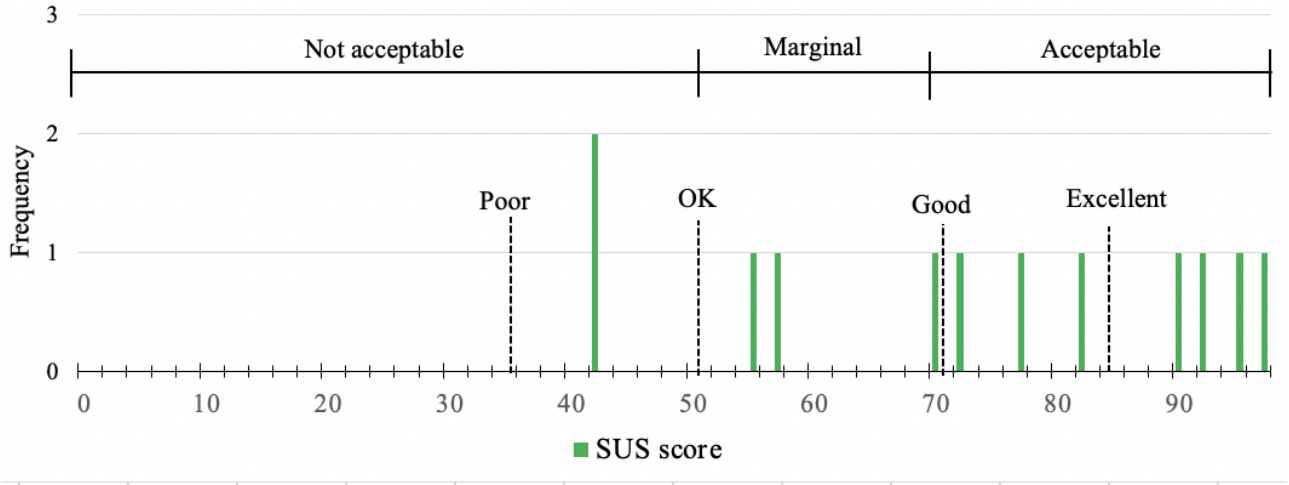
Usability of CARES

Overall, participants reported above-average system usability on the SUS, with a mean score of 72.92 (SD 19.77) and a range from 42.50 to 97.50. Most participants found CARES to be an acceptable (8/12, 67%) and good or excellent (7/12, 58%) system. Specifically, the current family caregivers receiving support reported above-average system usability, with a mean score of 76.94 (SD 19.03) and a range from 42.50 to 97.50. Caregiver peer supporters reported marginal usability, with a mean score of 60.83 (SD 20.21) and a range from 42.50 to 82.50 on the SUS. One current caregiver and one caregiver peer supporter rated CARES as below OK. The distribution of the acceptability and adjective ratings as indicated by Bangor et al [30] are shown in Figure 6.

On average, participants sent 27 (SD 6.88) messages, with a range from 15 to 36, to their assigned caregiver peer over the course of the 2 weeks. All participants engaged weekly with the app. On average, participants reviewed 44% of the library resources over the course of the intervention, with a range from 0% to 100%.

Figure 6 shows the distribution of participant and interventionist (N=12) responses to the SUS. The x-axis marks the individual SUS scores, and the y-axis marks the frequency of the scores. Acceptability ranges and adjective ratings are informed by the work by Bangor et al [30].

Figure 6. Distribution of results of the System Usability Scale (SUS) for study participants and interventionists.



Barriers to and Acceptability and Facilitators of Using CARES

Overview

Regarding the acceptability of the CARES app and intervention, 24 codes and 6 themes related to the acceptance of CARES were identified. The 6 themes were improving the CARES app and intervention, acceptability of the CARES app features and design, value of the caregiver-peer relationship, barriers and limitations of CARES, caregiver needs and preferences, and caregiver challenges. The themes are listed in Table 2.

XSL•FO
RenderX

Table 2. Themes from qualitative analysis of the semistructured interviews.

Theme	Description	Participant quotes
Improving the CARES[a] app and intervention	The participants and caregiver peer supporters provided input on how to improve the CARES app features and study protocol.	• "If you're dealing with some kind of messaging app...there's no point if there's no notification because nobody will think to go check it." [participant 5] • "One of my concerns was that a couple of the people didn't understand what they were supposed to be doing, or how to interact with the app. I think that there needs to be a little bit more explanation upfront before we start interacting with [participants]." [caregiver peer supporter 14]
Acceptability of the CARES app features and design	Most participants and caregiver peer supporters found the CARES app to be acceptable for providing support to caregivers of individuals with dementia.	• "I knew that right after, like the first couple of messages back and forth, I was like, this is a great idea. Because it's convenient. It's easy. It's, you know, nonjudgmental. It's just what you want from a support thing." [participant 7] • "I think that it's like a personal support group. That's what it is. And it's in your pocket, because it's on your phone and it's delivered in an app, you don't have to leave your home, you don't have to try to arrange coverage to somebody to sit with, you know, your loved one, so that you can sneak out for an hour and then worry the whole hour that you're out about what's going on at home." [caregiver peer supporter 13]
Value of the caregiver-peer relationship	Participants specifically highlighted their appreciation of the caregiver-peer relationship.	• "You really felt like you had somebody to reach out to in the times when things were really stressful and really felt overwhelming. It just was somebody that you could connect with that knew how you were feeling." [participant 7] • "...knowing that there's somebody out there who is thinking about me and my situation." [participant 10]
Barriers and limitations of CARES	This refers to challenges that users face in using the CARES app and delivering the CARES intervention.	• "The only barrier I see is, if somebody doesn't have access to an iPhone." [participant 8] • "I'm so busy doing the caregiver stuff, and all the other sort of managing, so that if I have any time to myself, I would want to be doing other things that, you know, that didn't involve caregiving. So I wouldn't be apt to wanting to take the time out of those personal times." [participant 10]
Caregiver needs and preferences	Most participants and caregiver peer supporters found that the CARES app met the needs and preferences of family caregivers of people with dementia.	• "I think that a lot of caregivers will love [CARES] too. You know, the doctors are doctors, they're doing the medical part of it. They don't even think about the emotional part that the caregiver is going through." [caregiver peer supporter 14] • "It's more I know that I have to take care of myself in order to be a better caregiver. And I can't do that if I'm not feeling good about myself. And yet, I didn't know how to do that...So I think I think [my peer supporter] was really good the way that she, she validated my feelings and, and was out there for me." [participant 6]
Caregiver challenges	Participants highlighted the overall challenges of caring for a family member with dementia.	• "My daughters were very concerned about me not getting out enough on my own." [participant 9] • "One of the biggest struggles that I have is finding people that understand what I'm going through...it's very difficult to find somebody that I could connect with that had been through what I was going through, and that I felt comfortable really voicing my feelings to. So I think the idea behind the app is like, great, actually because, I mean, it's like somebody that's always there that knows exactly what you're feeling and what you've been through." [participant 7]

[a]CARES: Caregiver Remote Education and Support.

Improving the CARES App and Intervention

The most prevalent theme was improving the CARES app and intervention. This theme comprised 3 subthemes: improving the study, improving the app features, and technological difficulties. Improving the CARES app and intervention constituted 26% of the themes discussed in the interviews. The participants and caregiver peer supporters provided input on how to improve the CARES app features and study protocol. For example, a few participants noted problems receiving notifications when their assigned peer sent them a message through the app. Participant 5 mentioned the following:

> *...if you're dealing with some kind of messaging app...there's no point if there's no notification because nobody will think to go check it.*

Some participants found aspects of the app difficult to maneuver. At times, the video chat would not connect properly, and the library resources would appear blank. Caregiver peer supporter 13 thought that the home page of the CARES app should clarify the content of the library resource feature to increase interaction with the materials contained within the library:

> *...you don't know that the resources are there, they're offering them to you, but you can't find them. It's like you're going on a board game, and you don't know where to get off.*

Other participants provided ideas on how to enhance the *app*. For example, caregiver peer supporter 15 suggested that the app include the option to communicate via telephone along with video chat and messaging as building rapport can be "very difficult to do with texting." Other participants suggested adding a support group–like feature where participants could connect with multiple peers rather than just 1. Participant 5 suggested the following:

> *...having kind of a group chat or a message board, where you could just be like, okay, just, you know,*

XSL•FO
RenderX

> *like, venting for a second or whatever, that could be really kind of an expansion from beyond just the one on one.*

The caregiver peer supporters also provided feedback on how to improve the peer support training. The caregiver peer supporters thought that clarification of the peer supporter and participant roles and expectations would have improved the participant-peer interaction and relationship. For example, caregiver peer supporter 14 stated the following:

> *...one of my concerns was that a couple of the people didn't understand what they were supposed to be doing, or how to interact with the app. I think that there needs to be a little bit more explanation upfront before we start interacting with* [participants]

Both participants and caregiver peer supporters recommended holding a training specifically on the features of the CARES app. Caregiver peer supporter 13 said the following:

> *I think that it would have also been beneficial to download the app, and then on part of the training, you walk through it with us, and we...just play with it...like hands on learning.*

The caregiver peer supporters suggested that the training should include additional practice using the CARES app and suggestions on how to initiate relationships with their assigned participants. Participants proposed that, in the future, a researcher should explain the features available on the CARES app and clarify their expectations for both the caregiver peer and the participant. Finally, participants recommended that future studies should match peers with participants based on their relationship to the individual with dementia they are caring for.

Acceptability of the CARES App Features and Design

The second theme, acceptability of the CARES *app* features and design, constituted 25% of the themes discussed in the interviews. Overall, most participants and caregiver peer supporters found the CARES app to be acceptable for providing support to caregivers of individuals with dementia. All participants agreed that the main purpose of the app was to connect caregivers with peers with a similar lived experience. Most participants interacted most with the messaging feature. Participant 5 mentioned the following:

> *I think the main point, or the main feature, is the connection to peers.*

Participants and caregiver peer supporters emphasized the convenience and accessibility of the CARES app. For example, participant 7 said the following:

> *I knew that right after, like the first couple of messages back and forth, I was like, this is a great idea. Because it's convenient. It's easy. It's, you know, non-judgmental. It's just what you want from a support thing. And it's also sort of on your own terms, though, because you get a message. So like, if you didn't want to respond instantly, you can kind of gather your thoughts, and you have time to respond, unlike a regular back and forth support group where if somebody asked me a question, I kind of have to have an answer...So. I think this was better, because I had a few minutes to really think through what she was asking me...and I had a minute to kind of gather my thoughts, and then I could just type it back.*

Participant 5 mentioned that, in contrast to in-person support groups where caregivers may "struggle with getting away even for an hour out of an apartment or their house for an hour," they "like the fact that they can [use CARES] over an iPad or an iPhone...I think that definitely makes it more accessible and easy." Caregiver peer supporters agreed that the CARES app was an appropriate tool for current caregivers of people with dementia. Caregiver peer supporter 13 shared the following:

> *I think that it's like a personal support group. That's what it is. And it's in your pocket, because it's on your phone and it's delivered in an app, you don't have to leave your home, you don't have to try to arrange coverage to somebody to sit with, you know, your loved one, so that you can sneak out for an hour and then worry the whole hour that you're out about what's going on at home.*

However, while all participants found the messaging feature of the CARES app beneficial, only some of the participants used the library resources (specifically the mindfulness, values, and negative thinking information). The "goals" and "wellness" features of the app were the least used by participants.

Value of the Caregiver-Peer Relationship

The third theme, the value of the caregiver-peer relationship, constituted 15% of the themes discussed in the interviews. Participants specifically highlighted their appreciation of the caregiver-peer relationship. Most participants emphasized that the purpose of the CARES app was the caregiver-peer connection. Participants found that their assigned caregiver peer supporters were knowledgeable, understanding, and validating. Participant 7 mentioned the following:

> *...you really felt like you had somebody to reach out to in the times when things were really stressful and really felt overwhelming. It just was somebody that you could connect with that knew how you were feeling.*

They also shared the following:

> *...it was also nice, because the person that I was connecting with chose to be connected with somebody. So it wasn't like you felt like you were burdening somebody else with your feelings.*

The caregiver peer supporters also found value in the caregiver-peer relationship. Caregiver peer supporter 13 said the following:

> *I think it's a great way to connect with caregivers. And it's easy because you can just type a message and somebody picks up that message at that point, and it's like having a support system at your fingertips...So I think when it works correctly, it would be a great effective tool for caregivers...because I found that caregivers in general, you know, they're*

XSL•FO
RenderX

> *not healthcare professionals. And they're expected to be in that role as part of the healthcare team. And they don't even know like, who to call, like, do I call you when my husband is running a fever and needs care?...I feel for these people that are just like plunked down in this role. And they don't have anyone. Even [one participant] said, I feel great, just knowing that you're there. And you provided this emotional support for me. And it's nice to know that you can reach out to someone that gets it. And that's exactly what it is...You know, it's a lifeline for people.*

Participants and caregivers found that participants appreciated the intervention "knowing that there's somebody out there who is thinking about me and my situation" (participant 10). The caregiver-peer connection was central to the CARES intervention.

Barriers and Limitations of CARES

The fourth theme, barriers and limitations of CARES, constituted 12% of the themes discussed in the interviews. Participants highlighted 2 main barriers and limitations: access to technology and time constraints. First, participants recognized that some caregivers may not have access to a smartphone or tablet or may not have adequate internet connection. Participant 8 mentioned that "the only barrier I see is, if somebody doesn't have access to an iPhone." Other participants noted that older adults may not be comfortable using technology. Time constraints and competing priorities were cited as other barriers that participants might face when using the CARES app. For example, participant 10 shared the following:

> *I'm so busy doing the caregiver stuff, and all the other sort of managing, so that if I have any time to myself, I would want to be doing other things that, you know, that didn't involve caregiving. So I wouldn't be apt to wanting to take the time out of those personal times.*

Finally, many participants and caregiver peer supporters thought that a 2-week period did not provide enough time to fully explore the features of the CARES app and peer relationship. For example, participant 7 said that "well, I didn't really check it out as much as I wanted to, because two weeks is not a lot of time."

Caregiver Needs and Preferences

The fifth theme, caregiver needs and preferences, constituted 11% of the themes discussed in the interviews. Overall, most participants and caregiver peer supporters found that the CARES app met the needs and preferences of family caregivers of people with dementia. Specifically, participants found that the CARES app and intervention provided social and emotional support. Participant 6 shared the following:

> *I think sometimes it's important for people who are caregivers to just be able to say how they're, how they're feeling.*

Caregiver peer supporter 14 mentioned the following:

> *I think that a lot of caregivers will love [CARES] too. You know, the doctors are doctors, they're doing the medical part of it. They don't even think about the emotional part that the caregiver is going through.*

For many participants, the CARES app and intervention provided individualized care and support. For example, participant 6 shared the following:

> *...it's more I know that I have to take care of myself in order to be a better caregiver. And I can't do that if I'm not feeling good about myself. And yet, I didn't know how to do that. Because I'm so tied up, so wrapped up in this. So I think I think she was really good the way that she, she validated my feelings and, and was out there for me.*

However, participants also stressed that caregiver peer supporters should have a general understanding of individuals' backgrounds and access to resources when providing support. For example, participant 5 mentioned the following:

> *...there's a wide range of resources that people have. I've seen this in the caregiver group I am a part of. Some people have planned well, and can afford help and support and some people don't have that, and they have no family around...The support person...should have an awareness of saying...you need to just hire somebody to come in for an hour every day.*

Caregiver peer supporters need to have an awareness of participants' available resources and priorities when providing support. Finally, most caregiver peer supporters also felt that CARES met their needs and preferences as former caregivers. For example, caregiver peer supporter 13 shared the following:

> *I always feel purposeful when giving back. That's most of the reason that I do coaching...I always feel a feeling of purpose. And there's a lot of emotional support and a feeling of gratification that comes from giving that emotional support, because you have the lived experience that you can share with these other caregivers. And if you don't have that experience, you don't get it like you can be in that role, but you don't truly get what they're going through.*

Caregiver peer supporters felt a sense of purpose while delivering the CARES intervention.

Caregiver Challenges

Finally, caregiver challenges was an emerging theme that constituted 5% of the themes discussed in the interviews. In their interviews, the participants highlighted the challenges of caring for a family member with dementia. Topics included difficulty taking time for oneself, frustration, anxiety about the unknown and upcoming changes, guilt, and the inability to find people who understand their situation. For example, participant 9 mentioned the following:

> *...my daughters were very concerned about me not getting out enough on my own.*

Participant 6 said the following:

XSL•FO
RenderX

> *I just wish with this disease they could say, well, in six months you may experience this and another six months you may experience that.*

Participant 2 stated the following:

> *...there are certain days you just feel more on top of things than others.*

Some participants shared how the CARES app and intervention addressed their challenges:

> *But one of the biggest struggles that I have is finding people that understand what I'm going through. Because if you have not provided care for somebody with dementia, you really don't understand like you can try to understand you can have the knowledge. But if you don't have the experience, it's very difficult to find somebody that I could connect with that had been through what I was going through, and that I felt comfortable really voicing my feelings to. So I think the idea behind the app is like, great, actually because, I mean, it's like somebody that's always there that knows exactly what you're feeling and what you've been through.* [participant 7]

Others did not voice whether CARES attended to the specific challenges they faced as caregivers.

Preliminary Effectiveness of CARES

Participants' decreases in burden, strain, and stress levels were not significant. However, we were not powered to find significance; rather, the purpose of this study was feasibility and acceptability. Non–statistically significant improvements were observed in all measures. The results of the baseline and 2-week posttreatment assessment for the 9 participants who completed the CARES intervention are shown in Table 3.

Table 3. Changes in outcomes from before to after the field usability study (2 weeks) for study participants (N=9)[a].

Measure	Participants, n (%)	Pretest assessment, mean (SD)	Posttest assessment, mean (SD)	*t* test (*df*)	*P* value	Effect size (Cohen *d*)
ZBI-12[b]	9 (100)	22.44 (10.04)	21.44 (11.34)	1.25 (8)	.25	0.42
MCSI[c]	9 (100)	11.22 (6.62)	11.00 (7.44)	0.41 (8)	.70	0.13
CSAQ[d] total	9 (100)	6.56 (3.97)	5.78 (4.12)	1.31 (8)	.23	0.44
CSAQ stress	9 (100)	5.28 (2.71)	4.89 (2.84)	1.00 (8)	.35	0.33

[a]A 2-tailed paired *t* test was used to assess statistical significance.

[b]ZBI-12: Zarit Burden Interview–Short Form.

[c]MCSI: Modified Caregiver Strain Index.

[d]CSAQ: Caregiver Self-Assessment Questionnaire. The CSAQ total is the average number of "yes" responses on the CSAQ, and the CSAQ stress score is the mean score on question 17 of the CSAQ, which asked participants to rate their level of stress on a scale from 1 to 10.

Discussion

Principal Findings

The purpose of this study was to evaluate the usability, acceptability, and preliminary effectiveness of the CARES app and intervention. The pilot study demonstrated that a 2-week peer-delivered and technology-supported mental health intervention (CARES) was acceptable for both former family caregivers of people with dementia who delivered peer support and current family caregivers of people with dementia who received peer support. Current caregivers reported above-average usability of CARES, and former caregivers reported marginal usability. The pilot study demonstrated that it is possible to train former caregivers in peer support and the delivery of CARES and integrated psychoeducation and mental health interventions using technology. CARES was associated with non–statistically significant improvements in burden, stress, and strain levels.

The usability of the CARES app was demonstrated using the SUS. Most caregivers found the CARES app to be an acceptable and good system with above-average usability. The CARES app allowed peer caregivers to provide individualized support and provided caregivers with access to evidence-based mental health resources on topics such as mindfulness and acceptance. The usability of CARES was also demonstrated through participants' capacity to use the smartphone and tablet app, completion of library resources on the app, and use of the messaging chat feature. Overall, most participants and caregiver peer supporters agreed that the CARES app and intervention were an acceptable tool to support family caregivers of people with dementia. However, participants and caregiver peer supporters identified areas in which the usability and acceptability of the app could be improved, and the caregiver peer supporters specifically reported marginal usability of the CARES app. Caregiver peer supporters may have reported below-average usability because of latency in messaging and other technological difficulties using the app. Future studies should examine the cause of differences in usability scores between current and former family caregivers and update the CARES app accordingly.

Participants suggested that improvement of technological features would strengthen the app's ability to achieve its purpose of connecting caregivers with peers. Some participants and peers faced technological difficulties while using the app. For example, at times, the app would not notify participants of new messages. This created a barrier in participants' and peers' ability to communicate efficiently and effectively. Participants also provided suggestions on how to improve the app features. For example, participants suggested adding a telephone feature to the app. Participants believed that adding a telephone feature would allow caregivers to communicate via the medium of their

XSL•FO
RenderX

preference and, therefore, increase their comfort level with technology and the caregiver-peer relationship. Participants also suggested adding a feature through which they could connect with more than one caregiver with a similar lived experience and suggested adding more caregiver-specific resources to the library resource page.

Despite technological limitations, most participants and peers found that the CARES app was an acceptable support intervention for family caregivers of people with dementia. Participants identified the caregiver-peer connection as the principal feature of the CARES app. Participants labeled the message-based support as convenient, easy-to-use, accessible, and individualized. The caregiver peer supporters were described as knowledgeable, understanding, validating, and supporting. Participants felt that the shared lived experience offered by the former caregivers better matched their needs and preferences for emotional and social support compared to professional medical and health care workers. On the other hand, the former family caregivers felt a sense of purpose and gratification while delivering the CARES intervention.

These findings suggest that technology- and peer-based interventions are usable and acceptable to family caregivers of people with dementia and that a smartphone app is a promising tool to support the mental and emotional health and well-being of family caregivers outside of an in-person or clinical setting. Task shifting informal caregiver digital mental health services to community members with lived experience has the potential to provide acceptable mental health interventions to family caregivers of people with dementia while addressing the current barriers and challenges with respect to accessing support. While it is estimated that nearly 153 million older adults will have dementia worldwide by 2050, mental health services for caregivers and their family members are limited due to an insufficient number of adequately trained geriatric mental health care providers [16,40]. Peer-delivered and technology-supported interventions have the potential to provide mental health services to family caregivers of people with dementia that are easily accessible and effective [36]. While caregiver psychosocial interventions have faced limitations due to time commitments, geographical location, requirements to meet in person, and failure to address the individualized needs of the caregiver, former and current family caregivers of people with dementia reported that the CARES app and intervention addressed the unique needs and experiences of consumers.

The results of the study support the hypothesis that former family caregivers of people with dementia have the knowledge and skillset to deliver trained peer support to current family caregivers. In previous studies, peers have been reported to be particularly effective at engaging participants in interventions. By sharing a lived experience, peers have the ability to develop alliances with participants and are viewed as more credible than traditional clinicians and providers [21]. With the use of technology-based messaging and support, participants were able to provide individualized support to caregivers at any time and location.

However, there are barriers and limitations to consider when using and delivering the CARES app and intervention. First, the CARES app is not accessible to caregivers who do not have a smartphone or adequate internet connection. In addition, time constraints may limit caregivers' ability to interact with the app and their assigned peer. Offering peer support to family caregivers of people with dementia may place stress on the interventionists. Caregiver peer supporters involved in the study should be offered mental health support while delivering the intervention. Finally, participants and peer supporters suggested that the abrupt ending of the 2-week field usability study may leave caregivers without the resources and support systems they came to rely on to manage well-being. Future studies should provide caregiver participants with additional caregiver support resources at the end of the CARES field usability study.

This study is not without limitations. First, one member of the research and intervention development team conducted a qualitative analysis of the interview data and identified codes and themes, which may have biased the results. In addition, member checking was not used to validate the findings and assess the accuracy of the qualitative results. Second, caregivers' input was not included in the initial development of the CARES app. Stakeholder engagement in the early stages of intervention development is essential to ensure that the modality and components are relevant to the community [41]. Future studies will incorporate caregivers' feedback to further improve and adapt the CARES app and intervention. Third, some participants and caregiver peer supporters experienced technological difficulties with the CARES app. For example, 22% (2/9) of the participants and 33% (1/3) of the caregiver peer supporters were unable to receive notifications and, therefore, had delayed responses to messages. Another caregiver peer supporter experienced challenges sending messages and, consequently, had to delay the start of their 2-week field usability study. These technological difficulties could have impacted the results of the field usability study and SUS. Fourth, the participant response rate was not tracked during the recruitment process. Future studies should track the response rate to improve recruitment procedures and decrease bias (eg, nonresponse bias). Fifth, data on the frequency of use of CARES app features such as video chat, goal setting, and wellness plans were not tracked across participants. This information will be helpful for further understanding the acceptability and usability of CARES. Sixth, the caregiver peer supporters were not assessed regarding whether they had learned the topics presented in the training, and the competence of the training was not assessed. Therefore, it is unknown whether the training sufficiently educated caregivers on the delivery of peer support. Seventh, the fidelity of caregiver peer support was not systematically evaluated through audio interactions; rather, message data determined high fidelity to the peer support model. Future research will assess message and audio interactions to determine fidelity to the caregiver peer support model of care through audio recordings on the app. Caregiver peer supporters' deviation from the training may have biased the results. Eighth, the participant eligibility criteria were broad and included all family caregivers of individuals with dementia aged ≥18 years. Criteria such as whether the caregiver lived with their care recipient and the number of hours spent caring for their relative with dementia could impact the acceptability, usability, and effectiveness of the CARES app and intervention. Future studies should

XSL•FO
RenderX

investigate the influence of differing caregiver roles and responsibilities on CARES. Finally, demographic information on factors such as education and income level was not collected, which may have affected the results and the perceived usability of the app.

While the aim of this study was to evaluate the usability of CARES and assess the acceptability and potential barriers to using the CARES technology, future studies would benefit from a larger sample size and a longer trial duration. In the qualitative interviews, participants shared that they wished they had a longer period to explore the app and the caregiver-peer relationship. While the length of the study and sample size were consistent with a field usability study, longer trials would allow participants to further assess the usability and acceptability of the app and whether it meets their needs and preferences as caregivers [42]. A longer trial would also more accurately reflect the length and fluctuation of the dementia caregiving experience. Future studies would also benefit from a randomized controlled design (eg, an intervention group with CARES and caregiver peer support vs a control group) and a control of covariables influencing the outcomes to evaluate the effectiveness of the CARES training, intervention, and app in reducing burden, strain, and stress in family caregivers of people with dementia. For example, future studies should examine the influence of age; relationship to the individual with dementia receiving care; stage of dementia; years providing care; severity of dementia symptoms; severity of the family caregiver's stress, strain, and burden levels; and the use of other caregiver support treatments and interventions on the effectiveness of the CARES app.

Future studies should also assign caregivers to peers based on dementia diagnosis or relationship to the individual with dementia. Matching caregivers with peers based on shared caregiver lived experiences may moderate the effectiveness of the CARES intervention. Future research would also benefit from a more diverse group of participants. Recruitment procedures should focus on recruiting a sample of participants representative of the demographics of the greater caregiver population. This includes recruiting more caregivers of color and caregivers who identify as male or nonbinary. Finally, future work should address the benefit of caregiver peer support for the family caregivers both delivering and receiving the mental health intervention. Caregiver peer support may have bidirectional effects. The caregivers providing support may see improvements in their mental health and well-being along with those of the participants they are supporting. As indicated by the study interview data, caregiver peer supporters felt a sense of purpose while delivering the intervention. Future studies should further assess the potential bidirectional influence of the CARES app and intervention.

Conclusions

This pilot study demonstrated that it is possible to train former family caregivers of people with dementia to use technology and deliver the CARES mental health intervention to current family caregivers of people with dementia. These findings provide preliminary evidence that a peer-delivered and technology-supported intervention designed to improve burden, stress, and strain levels is feasible and acceptable.

Acknowledgments

This study was funded by the Dartmouth Ethics Institute Sayles grant.

Data Availability

The data sets generated during and analyzed during this study are available from the corresponding author on reasonable request.

Conflicts of Interest

KLF is partners with Social Wellness LLC & Emissary Health, Inc.

Multimedia Appendix 1
Images of the Caregiver Remote Education and Support app features and distribution of the System Usability Scale participant results.
[DOCX File , 2987 KB - humanfactors_v11i1e41202_app1.docx]

References

1. Alzheimer's Association. 2020 Alzheimer's disease facts and figures. Alzheimers Dement 2020 Mar 10;16(3):391-460 [FREE Full text] [doi: 10.1002/alz.12068] [Medline: 32157811]
2. Alzheimer's disease facts and figures. Alzheimer's Association. 2024. URL: https://www.alz.org/media/documents/alzheimers-facts-and-figures.pdf [accessed 2024-03-17]
3. Gratão AC, Brigola AG, Ottaviani AC, Luchesi BM, Souza É, Rossetti ES, et al. Brief version of Zarit Burden Interview (ZBI) for burden assessment in older caregivers. Dement Neuropsychol 2019 Mar;13(1):122-129 [FREE Full text] [doi: 10.1590/1980-57642018dn13-010015] [Medline: 31073389]
4. Brodaty H, Donkin M. Family caregivers of people with dementia. Dialogues Clin Neurosci 2022 Apr 01;11(2):217-228. [doi: 10.31887/dcns.2009.11.2/hbrodaty]

XSL•FO
RenderX

5. Kim H, Chang M, Rose K, Kim S. Predictors of caregiver burden in caregivers of individuals with dementia. J Adv Nurs 2012 Apr 28;68(4):846-855. [doi: 10.1111/j.1365-2648.2011.05787.x] [Medline: 21793872]
6. Liu Z, Heffernan C, Tan J. Caregiver burden: a concept analysis. Int J Nurs Sci 2020 Oct 10;7(4):438-445 [FREE Full text] [doi: 10.1016/j.ijnss.2020.07.012] [Medline: 33195757]
7. Thornton M, Travis SS. Analysis of the reliability of the modified caregiver strain index. J Gerontol B Psychol Sci Soc Sci 2003 Mar 01;58(2):S127-S132. [doi: 10.1093/geronb/58.2.s127] [Medline: 12646602]
8. Walter E, Pinquart M. How effective are dementia caregiver interventions? An updated comprehensive meta-analysis. Gerontologist 2020 Nov 23;60(8):609-619. [doi: 10.1093/geront/gnz118] [Medline: 33226434]
9. Signe A, Elmståhl S. Psychosocial intervention for family caregivers of people with dementia reduces caregiver's burden: development and effect after 6 and 12 months. Scand J Caring Sci 2008 Mar;22(1):98-109. [doi: 10.1111/j.1471-6712.2007.00498.x] [Medline: 18269429]
10. Butler M, Gaugler JE, Talley KM, Abdi HI, Desai PJ, Duval S, et al. Care interventions for people living with dementia and their caregivers. Agency for Healthcare Research and Quality (AHRQ). 2020 Aug. URL: https://effectivehealthcare.ahrq.gov/products/care-interventions-pwd/report [accessed 2024-03-17]
11. Carter G, Monaghan C, Santin O. What is known from the existing literature about peer support interventions for carers of individuals living with dementia: a scoping review. Health Soc Care Community 2020 Jul 09;28(4):1134-1151. [doi: 10.1111/hsc.12944] [Medline: 31919974]
12. Ploeg J, Markle-Reid M, Valaitis R, McAiney C, Duggleby W, Bartholomew A, et al. Web-based interventions to improve mental health, general caregiving outcomes, and general health for informal caregivers of adults with chronic conditions living in the community: rapid evidence review. J Med Internet Res 2017 Jul 28;19(7):e263 [FREE Full text] [doi: 10.2196/jmir.7564] [Medline: 28754652]
13. Weems JA, Rhodes S, Powers JS. Dementia caregiver virtual support-an implementation evaluation of two pragmatic models during COVID-19. Geriatrics (Basel) 2021 Aug 19;6(3):80 [FREE Full text] [doi: 10.3390/geriatrics6030080] [Medline: 34449649]
14. Han J, Guo G, Hong L. Impact of professionally facilitated peer support for family carers of people with dementia in a WeChat virtual community. J Telemed Telecare 2022 Jan;28(1):68-76. [doi: 10.1177/1357633X20910830] [Medline: 32192386]
15. Marziali E, Garcia LJ. Dementia caregivers' responses to 2 Internet-based intervention programs. Am J Alzheimers Dis Other Demen 2011 Feb;26(1):36-43 [FREE Full text] [doi: 10.1177/1533317510387586] [Medline: 21282276]
16. Kunik ME, Mills WL, Amspoker AB, Cully JA, Kraus-Schuman C, Stanley M, et al. Expanding the geriatric mental health workforce through utilization of non-licensed providers. Aging Ment Health 2017 Sep 31;21(9):954-960 [FREE Full text] [doi: 10.1080/13607863.2016.1186150] [Medline: 27243369]
17. Barnett ML, Lau AS, Miranda J. Lay health worker involvement in evidence-based treatment delivery: a conceptual model to address disparities in care. Annu Rev Clin Psychol 2018 May 07;14(1):185-208 [FREE Full text] [doi: 10.1146/annurev-clinpsy-050817-084825] [Medline: 29401043]
18. Fortuna KL, Brooks JM, Umucu E, Walker R, Chow PI. Peer support: a human factor to enhance engagement in digital health behavior change interventions. J Technol Behav Sci 2019 Jun 15;4(2):152-161 [FREE Full text] [doi: 10.1007/s41347-019-00105-x] [Medline: 34337145]
19. Corey KL, McCurry MK, Sethares KA, Bourbonniere M, Hirschman KB, Meghani SH. Predictors of psychological distress and sleep quality in former family caregivers of people with dementia. Aging Ment Health 2020 Feb 27;24(2):233-241 [FREE Full text] [doi: 10.1080/13607863.2018.1531375] [Medline: 30588830]
20. Corey KL, McCurry MK. When caregiving ends: the experiences of former family caregivers of people with dementia. Gerontologist 2018 Mar 19;58(2):e87-e96. [doi: 10.1093/geront/gnw205] [Medline: 28064236]
21. Fortuna KL, DiMilia PR, Lohman MC, Bruce ML, Zubritsky CD, Halaby MR, et al. Feasibility, acceptability, and preliminary effectiveness of a peer-delivered and technology supported self-management intervention for older adults with serious mental illness. Psychiatr Q 2018 Jun 26;89(2):293-305 [FREE Full text] [doi: 10.1007/s11126-017-9534-7] [Medline: 28948424]
22. Solomon P. Peer support/peer provided services underlying processes, benefits, and critical ingredients. Psychiatr Rehabil J 2004;27(4):392-401. [doi: 10.2975/27.2004.392.401] [Medline: 15222150]
23. Zubin J, Spring B. Vulnerability--a new view of schizophrenia. J Abnorm Psychol 1977 Apr;86(2):103-126. [doi: 10.1037//0021-843x.86.2.103] [Medline: 858828]
24. Harris R. ACT Made Simple: An Easy-To-Read Primer on Acceptance and Commitment Therapy. Okland, CA: New Harbinger; 2009.
25. SAMHSA. Core competencies for peer workers in behavioral health services. SAMHSA. 2015. URL: https://www.samhsa.gov/sites/default/files/programs_campaigns/brss_tacs/core-competencies_508_12_13_18.pdf [accessed 2024-03-17]
26. Hayes SC, Pistorello J, Levin ME. Acceptance and commitment therapy as a unified model of behavior change. Couns Psychol 2012 Sep 26;40(7):976-1002. [doi: 10.1177/0011000012460836]

XSL•FO
RenderX

27. Rosenbaum S, Kantner L. Field usability testing: method, not compromise. In: Proceedings of the 2007 IEEE International Professional Communication Conference. 2007 Presented at: IPCC '07; October 1-3, 2007; Seattle, WA p. 1-7 URL: https://ieeexplore.ieee.org/document/4464060 [doi: 10.1109/ipcc.2007.4464060]
28. Faulkner L. Beyond the five-user assumption: benefits of increased sample sizes in usability testing. Behav Res Methods Instrum Comput 2003 Aug;35(3):379-383. [doi: 10.3758/BF03195514]
29. Lewis JR. The system usability scale: past, present, and future. Int J Hum Comput Interact 2018 Mar 30;34(7):577-590. [doi: 10.1080/10447318.2018.1455307]
30. Bangor A, Kortum PT, Miller JT. An empirical evaluation of the system usability scale. Int J Hum Comput Interact 2008 Jul 30;24(6):574-594. [doi: 10.1080/10447310802205776]
31. Epstein-Lubow G, Gaudiano BA, Hinckley M, Salloway S, Miller IW. Evidence for the validity of the American Medical Association's caregiver self-assessment questionnaire as a screening measure for depression. J Am Geriatr Soc 2010 Feb 27;58(2):387-388. [doi: 10.1111/j.1532-5415.2009.02701.x] [Medline: 20370867]
32. Technical assistance for users of the CFIR framework. The Consolidated Framework for Implementation Research. URL: https://cfirguide.org/ [accessed 2022-04-17]
33. Damschroder LJ, Aron DC, Keith RE, Kirsh SR, Alexander JA, Lowery JC. Fostering implementation of health services research findings into practice: a consolidated framework for advancing implementation science. Implement Sci 2009 Aug 07;4:50 [FREE Full text] [doi: 10.1186/1748-5908-4-50] [Medline: 19664226]
34. Lawson B, Sampalli T, Wood S, Warner G, Moorhouse P, Gibson R, et al. Evaluating the implementation and feasibility of a web-based tool to support timely identification and care for the frail population in primary healthcare settings. Int J Health Policy Manag 2017 Jul 01;6(7):377-382 [FREE Full text] [doi: 10.15171/ijhpm.2017.32] [Medline: 28812833]
35. Fortuna KL, Myers AL, Walsh D, Walker R, Mois G, Brooks JM. Strategies to increase peer support specialists' capacity to use digital technology in the era of COVID-19: pre-post study. JMIR Ment Health 2020 Jul 23;7(7):e20429 [FREE Full text] [doi: 10.2196/20429] [Medline: 32629424]
36. Fortuna KL, Naslund JA, LaCroix JM, Bianco CL, Brooks JM, Zisman-Ilani Y, et al. Digital peer support mental health interventions for people with a lived experience of a serious mental illness: systematic review. JMIR Ment Health 2020 Apr 03;7(4):e16460 [FREE Full text] [doi: 10.2196/16460] [Medline: 32243256]
37. Fortuna KL, Torous J, Depp CA, Jimenez DE, Areán PA, Walker R, et al. A future research agenda for digital geriatric mental healthcare. Am J Geriatr Psychiatry 2019 Nov;27(11):1277-1285 [FREE Full text] [doi: 10.1016/j.jagp.2019.05.013] [Medline: 31196619]
38. Kolb DA, Fry R. Towards an applied theory of experiential learning. In: Cooper C, editor. Theories of Group Process. Hoboken, NJ: John Wiley & Sons; 1975:33-57.
39. Watkins DC. Rapid and rigorous qualitative data analysis. Int J Qual Methods 2017 Jun 08;16(1):160940691771213. [doi: 10.1177/1609406917712131]
40. GBD 2019 Dementia Forecasting Collaborators. Estimation of the global prevalence of dementia in 2019 and forecasted prevalence in 2050: an analysis for the Global Burden of Disease Study 2019. Lancet Public Health 2022 Feb;7(2):e105-e125 [FREE Full text] [doi: 10.1016/S2468-2667(21)00249-8] [Medline: 34998485]
41. Petkovic J, Riddle A, Akl EA, Khabsa J, Lytvyn L, Atwere P, et al. Protocol for the development of guidance for stakeholder engagement in health and healthcare guideline development and implementation. Syst Rev 2020 Feb 01;9(1):21 [FREE Full text] [doi: 10.1186/s13643-020-1272-5] [Medline: 32007104]
42. Duh HB, Tan GC, Chen VH. Usability evaluation for mobile device: a comparison of laboratory and field tests. In: Proceedings of the 8th conference on Human-computer interaction with mobile devices and services. 2006 Presented at: MobileHCI '06; September 12-15, 2006; Helsinki, Finland p. 181-186 URL: https://dl.acm.org/doi/abs/10.1145/1152215.1152254 [doi: 10.1145/1152215.1152254]

Abbreviations

ACT: acceptance and commitment therapy
CARES: Caregiver Remote Education and Support
CFIR: Consolidated Framework for Implementation Research
CSAQ: Caregiver Self-Assessment Questionnaire
HIPAA: Health Insurance Portability and Accountability Act
PI: principal investigator
RADar: rigorous and accelerated data reduction
SUS: System Usability Scale

XSL•FO
RenderX

Edited by A Kushniruk; submitted 18.07.22; peer-reviewed by L Moo, F Puga, W Tang, H Perez, YH Tung; comments to author 08.11.22; revised version received 02.01.23; accepted 13.04.23; published 27.05.24.

Please cite as:
Collins-Pisano C, Leggett AN, Gambee D, Fortuna KL
Usability, Acceptability, and Preliminary Effectiveness of a Peer-Delivered and Technology-Supported Mental Health Intervention for Family Caregivers of People With Dementia: Field Usability Study
JMIR Hum Factors 2024;11:e41202
URL: https://humanfactors.jmir.org/2024/1/e41202
doi:10.2196/41202
PMID:38801660

©Caroline Collins-Pisano, Amanda N Leggett, David Gambee, Karen L Fortuna. Originally published in JMIR Human Factors (https://humanfactors.jmir.org), 27.05.2024. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on https://humanfactors.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Smartphone App Designed to Collect Health Information in Older Adults: Usability Study

Joanne M Murabito[1,2], MD, ScM; Jamie M Faro[3], PhD; Yuankai Zhang[4], MA; Angelo DeMalia[2], BA; Alexander Hamel[3], BA; Nakesha Agyapong[3], BA; Hongshan Liu[2], MA; Eric Schramm[5], PharmD; David D McManus[3,6], MD, ScM; Belinda Borrelli[7], PhD

[1]Section of General Internal Medicine, Department of Medicine, Boston University Chobanian & Avedisian School of Medicine, Boston, MA, United States

[2]Framingham Heart Study, Boston University, Framingham, MA, United States

[3]Department of Medicine, University of Massachusetts Chan Medical School, Worcester, MA, United States

[4]Department of Biostatistics, Boston University School of Public Health, Boston, MA, United States

[5]CareEvolution, Inc, Ann Arbour, MI, United States

[6]Cardiology Division, Department of Medicine, University of Massachusetts Chan Medical School, Worcester, MA, United States

[7]Center for Behavioral Science Research, Henry M. Goldman School of Dental Medicine, Boston University, Boston, MA, United States

Corresponding Author:
Joanne M Murabito, MD, ScM
Section of General Internal Medicine
Department of Medicine
Boston University Chobanian & Avedisian School of Medicine
72 E Concord St Suite L-516
Boston, MA, 02118
United States
Phone: 1 508 935 3500
Fax: 1 508 626 1262
Email: murabito@bu.edu

Abstract

Background: Studies evaluating the usability of mobile-phone assessments in older adults are limited.

Objective: This study aims to identify design-based barriers and facilitators to mobile app survey completion among 2 samples of older adults; those in the Framingham Heart Study and a more diverse sample from a hospital-based setting.

Methods: We used mixed methods to identify challenging and beneficial features of the mobile app in participants from the electronic Framingham Heart Study (n=15; mean age of 72 years; 6/15, 40% women; 15/15, 100% non-Hispanic and White) and among participants recruited from a hospital-based setting (n=15; mean age of 71 years; 7/15, 47% women; 3/15, 20% Hispanic; and 8/15, 53% non-White). A variety of app-based measures with different response formats were tested, including self-reported surveys, pictorial assessments (to indicate body pain sites), and cognitive testing tasks (eg, Trail Making Test and Stroop). Participants completed each measure using a think-aloud protocol, while being audio- and video-recorded with a qualitative interview conducted at the end of the session. Recordings were coded for participant usability errors by 2 pairs of coders. Participants completed the Mobile App Rating Scale to assess the app (response range 1=inadequate to 5=excellent).

Results: In electronic Framingham Heart Study participants, the average total Mobile App Rating Scale score was 7.6 (SD 1.1), with no significant differences in the hospital-based sample. In general, participants were pleased with the app and found it easy to use. A large minority had at least 1 navigational issue, most committed only once. Most older adults did not have difficulty completing the self-reported multiple-choice measures unless it included lengthy instructions but participants had usability issues with the Stroop and Trail Making Test.

Conclusions: Our methods and results help guide app development and app-based survey construction for older adults, while also giving consideration to sociodemographic differences.

(JMIR Hum Factors 2024;11:e56653) doi:10.2196/56653

XSL•FO
RenderX

KEYWORDS

mobile application surveys; mixed methods; electronic data collection; mHealth; mobile health; mobile application; mobile applications; app; apps; application; applications; digital health; digital technology; digital intervention; digital interventions; smartphone; smartphones; usability; usable; usableness; usefulness; utility; health information

Introduction

Adults aged 65 years and older are increasingly using tablets and smartphones and engaging with a range of technologies [1]. Technology use can reduce social isolation [2] and enhance communication with family members and health care providers, thereby, increasing well-being. Further, digital technologies have the potential to improve the health of older adults by facilitating symptom monitoring and self-care management as well as monitoring cognitive and mobility decline. However, older adults often lack confidence in their ability to use technology [3] and report needing assistance with new electronic devices and mobile apps [4]. They face unique challenges with technology including poor eyesight, hearing loss, fine motor skill and sensory limitations, and cognitive decline. These challenges make it essential to understand how technologies can be made more useful to older adults. Perceived value, usefulness, and impact on quality of life are important predictors of technology adoption in this age group [5,6]. In addition, a design that minimizes user frustration will enhance the use and lower the risk of leaving older users out of the technology revolution.

Older adults are often not well represented in user testing of technology [7] due to the restricted age range of research studies, physical or sensory impairments, or because technology studies may be less appealing to them. There are a growing number of smartphone apps that include opportunities for self-management of specific diseases and cognitive self-assessment but the quality and usability of the apps are often unknown especially among groups of older adults and adults from diverse race and ethnic populations [8,9]. In addition, health care providers and hospital systems are increasingly requesting that patients complete previsit health questionnaires electronically, which help with care efficiency and are preferred by providers [10,11]. However, older adults are less likely to access and use patient portals and may have unique needs influencing their use [12-14].

Usability information provides practical recommendations that can help increase patient responsiveness to electronically collected data. For example, studies that have evaluated the usability of mobile apps that assess fall risk demonstrated the importance of simple instructions and clear feedback such as a color change to indicate task completion [6,15,16] A mobile app designed for older adults with heart failure to report Patient Reported Outcomes Measurement Information System (PROMIS) measures demonstrated that these adults successfully returned the PROMIS data and an additional survey indicated high levels of usability [17].

We designed a smartphone app for use by community-dwelling older adults who are participants in the Framingham Heart Study (FHS). The smartphone app consists of surveys with different response formats (eg, multiple choice, pictorial, and tasks). The aim of this study is to identify design-based barriers and facilitators to mobile app navigation and survey completion through usability testing. We also sought to understand whether participant feedback differed depending on the response format of each measure. Because FHS participants were White, we enrolled a diverse sample of older adults at a second site to understand if any additional barriers to mobile app survey completion were observed given that digital literacy and preferences for using technologies can vary across older adults of different races and ethnic backgrounds [12,18]. Importantly, little is known about the usability of mobile apps for racially or ethnically diverse populations [8].

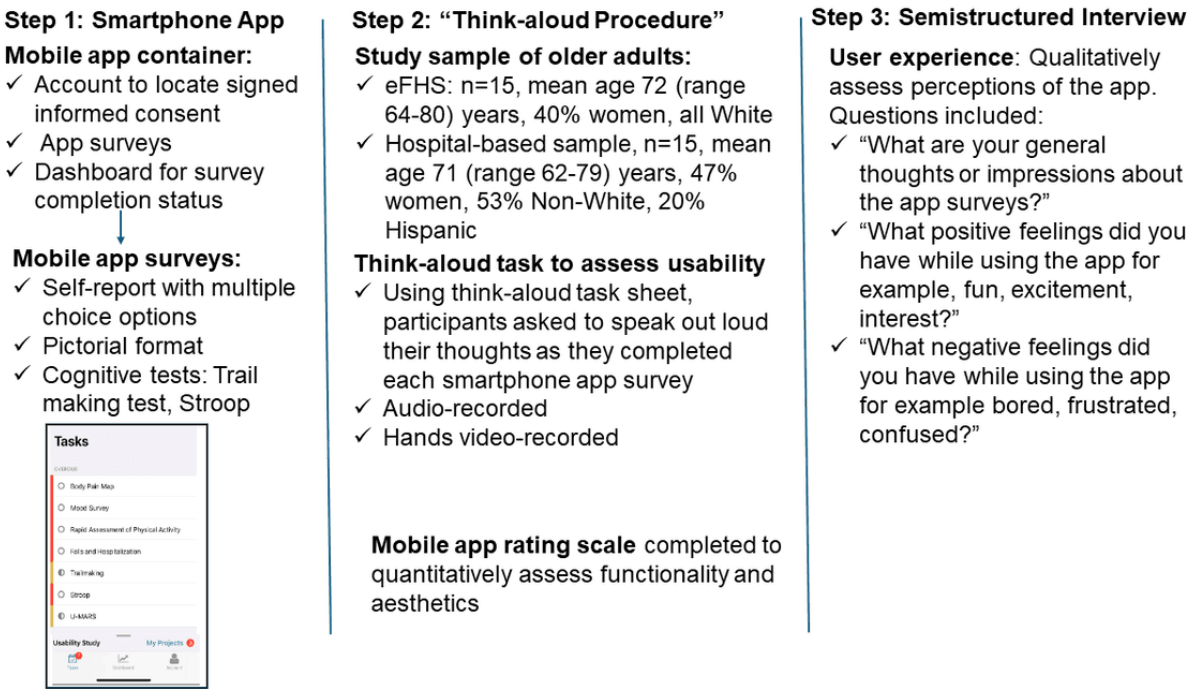
Methods

Study Design

The study used mixed methods to conduct 1 usability testing session followed by a postsession interview with enrolled participants (Figure 1). Usability testing methods included using the “think-aloud” protocol while conducting a series of surveys and tasks on the smartphone app. This was followed by a semistructured interview using an interview guide, to solicit information on barriers encountered in the session.

XSL•FO
RenderX

Figure 1. Overall study design.



Study Sample

The study sample was drawn from 2 sites: the FHS Offspring study and a hospital-based site. The FHS Offspring participants were recruited in 1971 and are invited back to the research center for examination every 6 to 8 years [19]. The tenth examination of the FHS Offspring participants occurred from 2019 to 2022 and included a mobile health component called the electronic Framingham Heart Study (eFHS). For this study, we enrolled eFHS participants who were English-speaking, owned a smartphone (iPhone or Android), attended Offspring exam 10 before the eFHS began, and enrolled in eFHS between July and September 2022. The eFHS research technician assisted the participants with registration, informed consent, and app download. Because most of the eFHS sample had iPhones and were less racially and ethnically diverse, we recruited participants from a second site who were not part of eFHS. The second site was a hospital in an urban area with a racially and ethnically diverse patient population. At the second site, inclusion criteria were age 60 years and older, English speaking, and able to attend a study session between December 2022 and April 2023. Participants were recruited through flyers placed in clinic waiting rooms and community centers, participation in prior research studies, and through patient registry lists. Non-White and Hispanic/Latino patients were oversampled from patient registry lists. As with eFHS, the research technician assisted the participants with registration, informed consent, and app download on a study iPhone. Participants at both sites were using the app for the first time. We enrolled 15 participants at each study site to ensure the representation of men and women, iPhone and Android users, and older people below and above 75 years. In addition, the number of unique challenges identified with the study design proposed appears to asymptote 15 participants [20]. While eFHS participants were not compensated for their time in the study (because FHS participants have not been compensated for participation in the parent study), participants from the hospital-based cohort were provided a US $100 card for participating upon completion of the session.

Ethical Considerations

The institutional review board at Boston University Medical Campus approved the eFHS study (H-36586) and this study (H-42659). The institutional review board at the University of Massachusetts Chan Medical School approved the hospital-based study (approval number 00000567).

Measures

Study App

A mobile app hosted a compendium of different types of survey assessments and tasks that users could click on to complete. CareEvolution's MyDataHelps Designer platform was used to build the smartphone app surveys for iPhone (iOS 10.0 or higher) or Android (version 7.0 or higher) devices. The MyDataHelps mobile app container includes an account where participants can locate their signed consent form, tasks (app surveys), and a dashboard. The dashboard was created in the app to provide the participant with survey completion status and encouragement with a thank you message. Investigators and CareEvolution industry partners internally tested the app surveys and tasks with attention to consistency inspection, and user-centered design principles to ensure clear instructions, easy navigation, and simple words and sentences [21].

Because the goal of the study was to assess usability, we included a variety of app-based surveys and tasks with different response formats (Multimedia Appendix 1). First, we tested several self-report surveys with multiple choice options including (1) the short form of the PROMIS measure of mood (anxiety and depression, 8-items) [22] and cognitive function

XSL•FO
RenderX

(4-items) [23] that included 5 response options that ranged from "never" to "always" for the mood assessment and "not at all" to "very much" for the cognitive function assessment; (2) falls and hospitalizations with yes/no response options and a calendar wheel to ascertain the date of a fall occurrence; and (3) rapid assessment of physical activity [24] to report the frequency of physical activity and the level of intensity (light, moderate, and vigorous), assessing frequency, intensity, and time, with yes/no response choices. The hospital-based site did not complete the PROMIS measure of mood. Second, we tested a pictorial format measure to collect data on chronic pain, based on a modified version of the Michigan body pain map measure [25]. Participants were shown an outline of a human body and were asked to click on the different places of the body where they are currently experiencing chronic pain, defined as pain lasting 3 months or longer. The app first displays the image of the front of the body and next the back of the body. If the participant is not experiencing chronic pain, a box was provided so that the user can indicate "no chronic pain." The third and final smartphone tasks tested were 2 commonly used cognitive tests. The Trail Making Test [26,27] is a timed assessment that requires the user to consecutively tap dots in alternating order between numbers and letters by first tapping the number "1" followed by tapping the letter "A" and then "2" followed by "B" until the user reaches the number "7." Correct answers result in the appearance of lines between the dots, resulting in a "trail." Finally, we assessed the Stroop [28] on the smartphone that requires the user to complete a series of 4 increasingly more difficult tasks responding as quickly and accurately as possible to changes in color and instructions. Because the Stroop requires the user to be able to see colors (yellow, green, red, and blue), persons with color blindness are not eligible for this task. A practice session was provided for each set of the 4 tasks with the ability to repeat the practice session should the user desire to do so [29]. At the end of the testing session, participants completed the Mobile App Rating Scale (MARS) on the smartphone to assess app functionality and aesthetics including ease of use, navigation, visual appeal, performance, graphics, and layout [30]. Items were rated by the participant using a 5-point Likert scale from 1=inadequate to 5=excellent.

Demographic data were not collected within the study app. For eFHS participants, these data were collected as part of Offspring exam 10. Sociodemographics such as age, gender, and employment were assessed via the research study coordinators and entered into a secure web-based software platform, REDCap (Research Electronic Data Capture; developed initially at Vanderbilt University, now collaborative support from the REDCap consortium) at the hospital-based site.

Procedure

Study sessions were conducted in person at the FHS Research Center. eFHS participants were asked to complete the above assessments using their own smartphones, but hospital-based participants used a study smartphone (iPhone 7). In addition, they were also asked to navigate to different areas of the MyDataHelps app container (account, tasks, and dashboard). While doing these tasks, participants were asked to "think your thoughts" out loud, including feelings (positive like "fun" and negative like "frustrating"). Prior to beginning the usability testing, the research technician demonstrated the "think-aloud" procedure using the text app on a smartphone and then asked the participant to demonstrate the think-aloud procedure using the same app verbalizing every movement, feeling, and decision. Once the participant was ready to begin, the research technician encouraged the participant to use the think-aloud task sheet (Multimedia Appendix 1) that included a list of app-based tasks. Participants were asked to speak out loud their thoughts, feelings, and actions as they completed each task. The participant was audio-recorded, and the participant's hands were video-recorded throughout the think-aloud procedure. The participant also completed the MARS [30] on the smartphone after the think-aloud procedure. The technician was present to audio- and video-record the think-aloud procedure and to encourage participants to speak their thoughts out loud. The technician was explicitly trained not to assist the participant with the app unless the participant was irrevocably stuck.

After completion of the think-aloud procedure, the research technician conducted a 15-20–minute interview using open-ended questions and reflective listening to obtain participant feedback on their experiences using the app. Interview questions are available in the Multimedia Appendix 1 and include questions such as "What are your general thoughts or impressions about the app surveys?"; "What positive feelings did you have while using the app for example, fun, excitement, interest?"; and "What negative feelings did you have while using the app for example bored, frustrated, confused?" In addition, the interviewer asked the participants to what extent family, friends, and people of their own age would be able to use and enjoy the app and their thoughts on how to ensure the app would be acceptable to people of different cultures. The research technicians (AD, NA, and AH) are coauthors of this work.

Research Technician Training

In total, 4 interviewers were trained (by BB and JF), 3 were bachelor level and 1 was PhD level. Training consisted of learning the think-aloud procedure and also how to conduct the post think-aloud interview. Building rapport and communication skills (open-ended questions and reflective listening) were part of the training. Training included didactics and role plays and trainees were required to complete 3 sessions with "mock" participants, supervised by 1 or both trainers before being cleared to do the protocol with study participants. Feedback on study participants was provided to research technicians on an ongoing basis, by viewing the audio and video tapes together. The audiotaped portion of the think-aloud procedure and interview was professionally transcribed (Daily Transcription).

Process of Coding the Sessions

Investigators developed a coding sheet and accompanying coding manual for use by teams of coders when coding the video- and audio-recordings of the think-aloud procedure and postinterviews. The coding sheet included general items (eg, navigation between surveys, tapping in the wrong area to advance to the next task, and unclear instructions for surveys) as well as assessment-specific variables (clarity of instruction, understanding concepts, navigating within a survey, and "look and feel" eg, font size, line spacing, and color). For training

XSL•FO
RenderX

purposes, all coders reviewed 3 participant recordings together and resolved any coding discrepancies before separating into the 2 coding teams to code in pairs. In order to maintain coding reliability over time, the 2 teams (team one: JF, JM; team 2: BB, DDM) independently coded the same participant recordings on 5 occasions throughout the coding process and came together to review the coding sheets for any discrepancies across teams to ensure all coders were following the coding guidelines. The average percent agreement between coders ranged from 80.5% to 98% across the 5 recordings that were coded in common by the 2 teams.

Statistical Analysis

Descriptive statistics of the 2 study samples used mean and SD for continuous variables, and numbers and percentages for nominal variables. For the comparison of continuous variables, 2-tailed *t* tests were applied, and chi-square tests or Fisher exact tests were used to compare nominal variables between the samples. The percent agreement between coding teams was calculated by the ratio of the number of discrepancies divided by the number of items in the coding sheet. For the MARS, all items used a 5-point Likert scale and the mean score for each domain (functionality and aesthetics) was calculated separately and an overall MARS score was computed for each of the 2 study samples. In addition, we calculated the mean score of each item within a domain (eg, layout, graphics, and visual appeal).

Results

Study Sample

In eFHS, 15 participants signed informed consent (mean age of 72, SD 4.2, range 64-80 years; 6/15, 40% were women; 1/15, 100% non-Hispanic, White). All eFHS participants owned a smartphone with 9 of 15 (60%) of the eFHS study sample participants owning an iPhone (Table 1). In the hospital-based site, 77 were contacted, 19 declined to participate, 3 deferred enrollment to a later date, and 15 participants signed informed consent (mean age 70.6, SD 6.2, range 62-79 years; 7/15, 47% women; 3/15, 20% Hispanic/Latino; and 8/15, 53% non-White). In the hospital-based sample, 1 participant did not own a smartphone, and among smartphone owners, 6 of 14 (42.9%) owned an iPhone. More than half of the participants at both sites had a college education or advanced degree. While nearly 90% (13 of 15 participants) of eFHS participants reported their health to be very good to excellent, only one-third (5 of 15 participants) of the participants at the hospital-based site did (*P*=.003).

Table 1. Participant characteristics.

Characteristics	eFHS[a] (n=15)	Hospital-based sample (n=15)	*P* value
Age (years)			.48
Mean (SD)	72 (4.2)	70.6 (6.2)	
Range	64-80	62-79	
Women, n (%)	6 (40)	7 (46.7)	.71
Non-White, n (%)	0 (0)	8 (53.3)	.002
Hispanic, n (%)	0 (0)	3 (20)	.22
Smartphone owner, n (%)[b]			
iPhone	9 (60)	6 (42.9)	.36
Android	6 (40)	8 (53.3)	.36
Bachelor degree and higher, n (%)	10 (66.7)	9 (60)	.71
Marital status (married, living as married), n (%)	12 (80)	8 (53.3)	.12
Income <US $55,000, n (%)[c]	4 (31)	5 (33)	.89
Subjective health, very good or excellent, n (%)	13 (86.7)	5 (33.3)	.003

[a]eFHS: electronic Framingham Heart Study.

[b]All offspring participants owned a smartphone (iPhone or Android); 1 hospital-based participant did not own a smartphone.

[c]Income of 2 participants in eFHS sample was unknown.

In the eFHS sample, the average length of time of the think-aloud procedure was 25.5 (range 12.4-44.0) minutes and the postprocedure interview time on average was 18.7 (range 12.4-20.9) minutes. Two participants declined the interview (think-aloud times were 28.5 and 33.35 minutes, respectively). At the hospital-based site, the average length of time of the think-aloud procedure was approximately 11 minutes longer (mean 36.5, range 24.1-55.1 minutes) while the postprocedure interview time was similar to eFHS with an average of 18.5 (range 10.6-35.1) minutes. No participant at the hospital-based site declined the post think-aloud interview.

Barriers to App Use Identified During Think-Aloud and Postprocedure Interview

Table 2 presents the themes identified from the think-aloud task and postprocedure interview along with sample participant quotes.

XSL•FO
RenderX

Table 2. Barriers and facilitators identified during the think-aloud and postprocedure interview.

Domain or survey type	Theme	Sample quotes
General structure, satisfaction	• Information about security of app • Use to raise awareness of health • Everyone could learn from it	• "The more I do, the easier it gets" • "I just think it's gonna help me monitor my health and go from there. I like technology that helps me and doesn't just amuse me or keep going and if this helps me stay healthy, stay fit, it makes sense."
Ease of use	• Easy, simple, fun • Questions easy to answer • Confusion related to not paying attention and not reading the app instructions carefully enough • Font color yellow difficult to see • Font could be bigger	• "The app spells everything perfectly clear after you read it a couple times to get it" • "I wasn't really paying attention to what I was reading. That was mostly my problem."
Navigation	• Navigating within and between surveys was easy • More detail on getting back to the home screen and using "next," "cancel," and "done"	• "The more I did it, the more I was able to figure out how to get from one place to another." • "You might want to give a little more detail about some of the sections, like what happens when you cancel, what happens when you hit done."
Multiple-choice survey formats	• Surveys were interesting and simple except physical activity survey with lengthy definitions and confusing response choices	• "It was easy the questions were simple and it was easy to find the answers."
Pictorial format	• Body pain map did not include all areas that can be a real problem • Checkbox worked differently than response choices for other surveys • Hard to find the "no pain" box—small font	• "Well, on the pain thing, I think they should have something near the anus. Because that can be a real headache. And then, of course, for women, it would also include, uh, the uterus."
Cognitive tasks: Trail Making Test and Stroop	• Stroop was confusing • Difficulty understanding practice session versus testing session • Confusion over 4 increasingly challenging sets of tests within the Stroop task • TMT[a] had difficulties with instructions and navigation but the eFHS sample did not	• "The all underling thing was a little confusing to me. I think because I was trying to get through it quick. I lost track of whether underline the word or is it the color or is it this color or the word?" • "The trail-making one I liked"
Friends or family	• Fear of technology; • Little interest in learning how to use technology • Learning curve for older people	• "I think they would find it very useful I think it's very useful, just for like, the cognitive part of it."
Different cultural backgrounds	• Need the app to be available in different languages	• "I come from the Indian community, and we place a lot of value on education, you know, I think this is the kind of thing that, you know—I would—we do—I would like to be quizzed, and this is a quiz." • "For example, I came from Burma. Burma used to be very underdeveloped country, and, also, still lots of problem. But they are very good at technology."

[a]TMT: Trail Making Test.

General Structure of the App

In general, participants were pleased with the general structure of the app. However, a few participants had not used an app previously and they said it was "not intuitive" but "once you run through it a few times it isn't a problem." Participants generally thought that the app had the right amount of surveys (and that additional surveys would make them bored). Participants suggested that a progress indicator be inserted to let users know the status of survey completion. One participant also raised the importance of conveying information regarding the security of the app.

Ease of Use

Overall, participants found the app "easy to use," "simple," and "fun." One participant said that it looked and functioned similarly to apps he was already familiar with, such as a social security app. A small minority of participants, however, wanted increased font size and more user-friendly colors. One participant was frustrated with the functionality of the app, but he blamed it on using an old phone. There were some

XSL•FO
RenderX

suggestions for modifications, such as tailoring the app by age group and disability, so it is not so "one size fits all." Another participant suggested removing the free text space (which was needed for 1 question) or adding a digital keyboard so it is more intuitive that the person is supposed to type. For example, upon completion of the Stroop task, there was a query ("did you encounter any issues during the task?") that permitted the participant to freely type in a response.

Navigation

A large minority of participants reported at least 1 navigational issue, including navigating within and between surveys and also clicking the incorrect area in the app to navigate to the next screen (eg, navigating to "done"). The vast majority of these errors were committed only once. Participants suggested making the buttons look more like buttons or having them flash when you are supposed to press the button, "...making it extremely obvious what you are supposed to hit." Participants also suggested that more details should be added, such as what will happen after you hit a button (eg, after "cancel" or "done" is pressed), and more information about what the "icons at the bottom of the app" mean (eg, dashboard). A few participants mentioned that the app should avoid having to scroll to see more information, and instead put all of the information on one page if possible or continue to the next page.

Response Choice Formats

Multiple-Choice Formats

Participants reported that most of the multiple-choice surveys were easy and interesting to complete, with only a few sporadic usability errors by a few participants. For the most part, participants reported that the instructions were easy to understand, with one exception, which was "The Rapid Assessment of Physical Activity" survey, which included lengthy definitions of physical activity levels (eg, mild, moderate, and vigorous activity) which were needed to answer the subsequent questions. Participants were confused by the content and format of the definitions. With regard to the latter, both physical activity intensity level and physical activity frequency in a single question (eg, "I do some light or moderate physical activities but not every week," yes or no).

Pictorial Format

For the pictorial format assessment, (modified version of the Michigan Body Pain map), a small minority of participants noted the checkbox for response worked differently than the other surveys; instead of seeing a checkmark in the response box, the response box if selected was highlighted in color and was confusing for some. Some participants also noted that the map did not include all body areas where pain was present. For some participants, the "no chronic pain" response box was difficult to find due to the small font, and the small font used to label the body parts was also difficult for several participants (but necessary to eliminate scrolling to see the entire body outline). There were no participants who reported difficulty understanding the instructions or concepts. Only a couple of participants did not realize that they had to press "next" to go to the next page to view the back of the body to indicate pain areas there. One participant said that bowels and reproductive areas give people his age a lot of issues and that these areas should be added to the body pain map.

Cognitive Testing Formats

In terms of the formats used for cognitive testing, most older participants had some difficulty with the process of completing these measures. In terms of the Stroop, some participants had difficulty understanding that they were in the practice session versus the test session. Participants also thought that they needed to do the testing quickly and were confused if they did not read the instructions for each of the 4 test sets, as each had different instructions. Most participants did not take advantage of the opportunity to repeat the practice session when asked by the app, even when they were confused by the task. The yellow font color for the Stroop was also problematic for a few participants. One participant suggested that voice and animation be used to explain the instructions for the Stroop. For the Trail Making Test, the eFHS sample did not report any difficulties with instructions or navigating within the task, and seemed to have a good understanding of the concepts included in the task. However, a large minority of the hospital-based sample had difficulties in all 3 of these areas. There were no difficulties reported on the look and feel of the task (eg, font, line spacing, and color).

Relevance for Friends, Family, and Different Cultures

Participants noted that they knew older people who engaged with technology and others who were not interested in the "electronic age," did not use a computer or smartphone, and had no interest in learning. One participant said that she had a friend aged 90 years or older who would easily be able to interact with the app and another friend of the same age who would have more difficulty. Participants noted older adults may have a fear of technology, be less confident using technology, and need assistance or a training session given that, for older people who have not used computers or technology, using the app would "be like a foreign language to them." Participants provided their thoughts on using the app in older adults of different cultural backgrounds. Many noted that the current version of the app is available only in the English language and would need translation to other languages. One participant was from a country that they felt embraced technology, and another participant reported her culture valued education and felt like the app included an educational-like component like a "quiz" which would be viewed positively by her culture.

Satisfaction With the App

The mean total MARS score (7.6, SD 1.1), mean functionality score (3.8, SD 0.6), and mean aesthetics score (3.8, SD 0.6) in the eFHS sample did not significantly differ from the hospital-based sample (Figure 2). With the exception of ease of use, the individual items of the functionality and aesthetics scores (performance, navigation, interactions such as taps, swipes, layout, graphics, and visual appeal) did not significantly differ between the 2 samples (Figure 3). Ease of use may have differed as participants at the hospital-based site used a study iPhone whereas eFHS participants used their own smartphone. The performance item was rated the highest with a mean of 4.5 in both the samples. The mean overall star rating was 3.5 (SD

XSL•FO
RenderX

0.7) in the eFHS sample and 3.7 (SD 0.96) in the hospital-based sample indicating participants rated the app above average. In addition, during the interview, participants noted that the app could be used to raise awareness of health and people could learn from it, even though that was not the original purpose of the app. Some participants liked the app because it enabled one to "express yourself" through the surveys. Finally, a few participants suggested making the app "more entertaining" by adding narration.

Figure 2. MARS scores by study sample: overall MARS functionality and aesthetics scores. MARS: Mobile App Rating Scale.

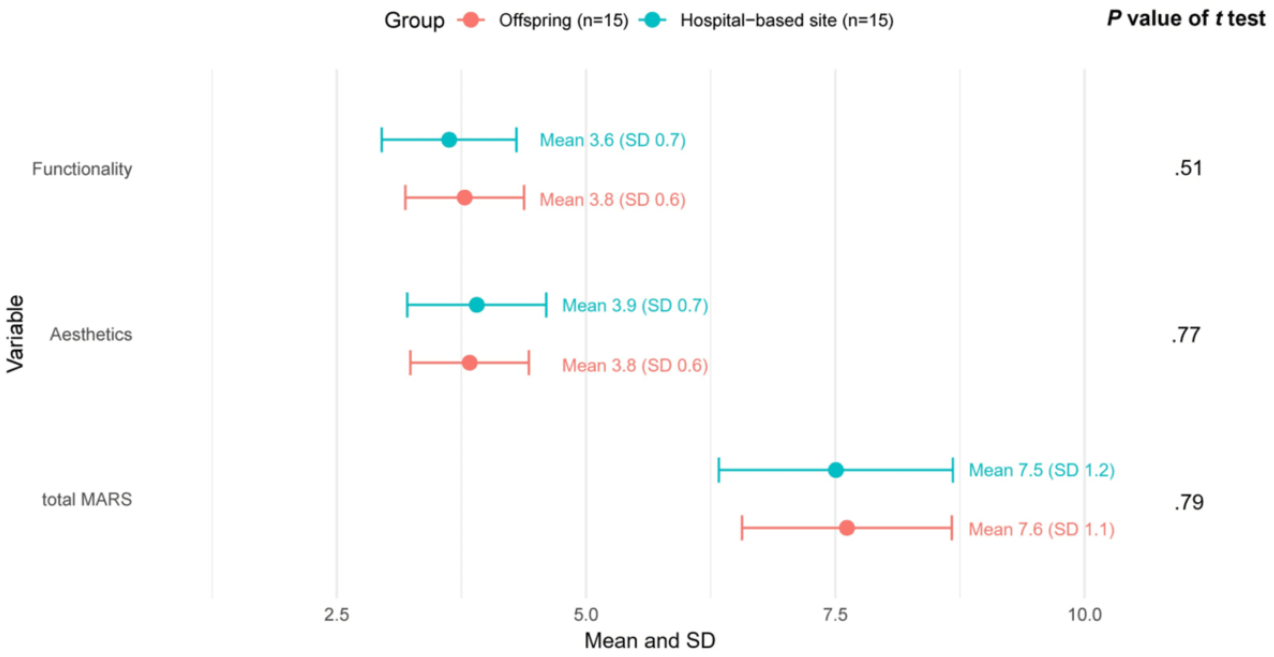
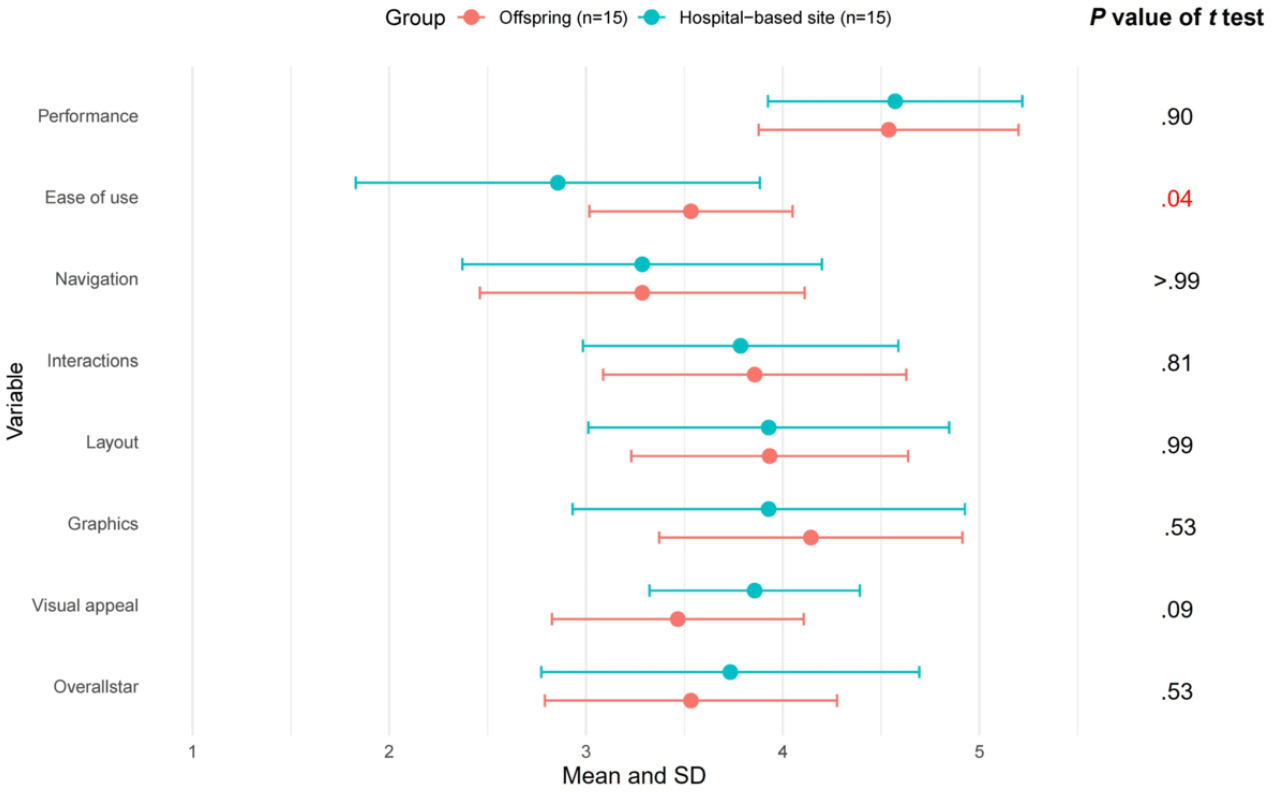


Figure 3. MARS (Mobile App Rating Scale) scores by study sample: individual items within the Functionality and Aesthetics domains.



XSL•FO
RenderX

Discussion

We tested the usability of a smartphone app designed to collect health information using complementary approaches in a community sample of older adult participants of the Framingham Offspring Study, and to understand generalizability we also tested older adults from a more diverse hospital-based sample. In general, participants liked the structure of the app and found the app was simple, fun, and easy to use. However, a large minority reported navigation issues that mostly occurred once with the ability to learn and figure out how to move within and between app-based surveys. A small minority of participants verbalized a preference for larger font sizes or more user-friendly colors. We observed that most older adults did not have difficulty with the multiple-choice app-based surveys unless the survey included lengthy instructions. Finally, most older adults experienced challenges with the app-based cognitive tasks especially the Stroop which required participants to read and understand a series of 4 increasingly more challenging tests with the Stroop task. Of importance, some participants noted that the app could be used to raise awareness of health and one could learn from it. Our observations confirm those of others that the involvement of older users can result in positive feelings among older adults, dispel stereotypes associated with older users, and the insights gained from older users can be used to enhance the quality of the design [31]. For example, our participants suggested using voice instructions and animated tutorials. To enhance usability, app designers and investigators should consider training that includes tutorials within the app provided by an older adult guide to boost confidence when designing smartphone apps for use by older adults.

This work has several implications for tailoring technology for older adult users. First, a guide within the app explaining the purpose of the app and highlighting key app functions including such functions as “next,” “back,” and “done” would enhance usability. Older adults are more likely to engage with technology that they perceive as useful [5]; therefore, having a clear understanding of the goals of technology is critical. Further, the addition of basic training in smartphone use in older adults less proficient with technology was associated with fewer errors and less cueing during a smartphone app–based health prevention program and may result in improved engagement [32]. In our sample, older adults who had not used an app before did not find it intuitive but after running through it a few times did not find it a problem. Our results support recently published app design guidelines for older adults advocating for initial training, if possible face-to-face, along with video instructions that are contextualized and provide step-by-step instructions to support older users [33]. Training may boost confidence and make the experience as frictionless as possible lowering the potential for abandonment. Second, streamlining and simplifying instructions may enhance understanding by inviting older adults to read them attentively. Participants in our study noted confusion that they attributed to not paying attention or related to the need to slow down and read directions more than once. Consistent with our observation, others have noted the need for clear and simple instructions when designing mHealth apps for older adults [6]. Some participants also requested features they enjoyed in their use of other apps. Gamification of functions, where possible, such as a flashing “done” button or a countdown to the start of the next task may improve engagement. Finally, consider voice narration and animated guides throughout the app surveys where features other than straightforward multiple-choice questions and responses will be encountered.

Our study may have important insights to help address the continued digital disparities observed in older adults. Older adults are increasingly using the internet and smartphones [34]; however, connection to the internet decreases across ages with nearly half of young adults almost constantly connected versus 8% of adults aged 65 years and older [4]. Other key digital health behaviors are also lower in a nationally representative sample of older adults including using health apps, using a digital device to track health or a health-related goal, and digitally communicating with a health care provider [35]. Digital technologies were lifesaving during the COVID-19 pandemic as the rapid transformation from in-person visits to televisits permitted access to health care in a setting that provided social distance and did not expose vulnerable older adults to the virus. Similarly, the ability to participate in digital interventions may provide several benefits, such as improved memory and independent living [36], physical functioning, physical activity [37], depression, and anxiety [38]. Including older adults in technology design and conducting usability testing may address digital health inequities by addressing digital health literacy and creating programs that are user-friendly to this population [39]. They may also improve implementation beyond pilot studies and achieve the needed sustainability of technology solutions [40] for chronic disease management and home care options for older adults and, at the same time, maybe one step in addressing digital disparities. We plan to use the smartphone app more widely in the Framingham cohorts as a tool to monitor health. We will be able to provide critical information on the characteristics of those who enroll and use the technology, as well as those who choose not to.

Our study had several strengths. There is no “best” method to assess usability [21]; therefore, we used both qualitative and quantitative methods. We tested older users with mean age of 70 years and older in both samples often not included in studies testing technology and included older adults from diverse race and ethnic backgrounds. The study sample included a range of older users. Participants without a smartphone or experience with app use and both iPhone and Android users were included. This strategy allowed us to uncover errors with the app beyond what would be observed with “regular” users and permit greater guidance in app redesign to benefit older users.

In addition, participants with health issues were included. Some older adults with health conditions or geriatric syndromes such as frailty have higher levels of nonuse of information communication technologies and more negative views on usefulness and usability [41]. Our study also had some limitations that merit comment. Participants at the hospital-based site used an iPhone only. This may have been a limitation if the participants were Android owners or had a different iPhone version; however, this is also a strength as we were able to include participants who were not smartphone owners. Our observations focus on the first interactions with the app-based

XSL•FO
RenderX

surveys. It is beyond the scope of the study to examine other aspects of use such as efficiency (how quickly the survey or task is completed once the design is learned) and memorability (how easy to use after a period of not using the app) that may have important implications to research study designs. Continued engagement with technology changes over time in older adults but the factors related to continued technology use are unclear and require further investigation [42]. Our study took place in Massachusetts, and therefore, may not be generalizable to other geographic areas. Participants with color blindness were not eligible for the Stroop task and were excluded from testing. Therefore, results may not be generalizable to this group of older adults. FHS participants who enroll in eFHS are healthier and have higher levels of education than participants who chose not to enroll.

Our study of a diverse sample of older adults testing several different smartphone app survey types and response formats provides a guide to investigators and clinicians that can be used for future app development and app-based survey construction for older adults. Many older users are able to interact with and enjoy technology. Further work to enhance engagement among older users and diminish digital disparities in this group is needed if the potential of technology to improve well-being, functioning, and health in older adults is to be realized.

Acknowledgments

This work was supported by the National Institutes of Health, National Institute of Aging (R01AG076757) to JMM. The Framingham Heart Study is supported by the National Heart, Lung, and Blood Institute at the National Institutes of Health (contracts N01-HC-25195, HHSN268201500001I, and 75N92019D00031).

Conflicts of Interest

DDM has received research support from Fitbit, Apple Inc, Bristol–Myers Squibb, Boehringer–Ingelheim, Pfizer, Flexcon, Samsung, Philips Healthcare, and Biotronik, and has received consultancy fees from Heart Rhythm Society, Bristol–Myers Squibb, Pfizer, Fitbit, Flexcon, Boston Biomedical Associates, VentureWell, Avania, and Rose Consulting. DDM also declares financial support for serving on the steering committee for the GUARD-AF study (NCT04126486) and the advisory committee for the Fitbit Heart study (NCT04176926). ES is an employee of CareEvolution, Inc, a health care technology company. The remaining authors have no conflicts of interest to declare.

Multimedia Appendix 1
Screenshots of app-based surveys and tasks, think- aloud task sheet, and post procedure interview questions.
[DOCX File , 3864 KB - humanfactors_v11i1e56653_app1.docx]

References

1. Faverio M. Share of those 65 and older who are tech users has grown in the past decade. Pew Research Center. 2022. URL: https://tinyurl.com/2fza68h7 [accessed 2022-04-11]
2. Umoh ME, Prichett L, Boyd CM, Cudjoe TKM. Impact of technology on social isolation: longitudinal analysis from the National Health Aging Trends Study. J Am Geriatr Soc 2023;71(4):1117-1123. [doi: 10.1111/jgs.18179] [Medline: 36519748]
3. Lee CC, Czaja SJ, Moxley JH, Sharit J, Boot WR, Charness N, et al. Attitudes toward computers across adulthood from 1994 to 2013. Gerontologist 2019;59(1):22-33 [FREE Full text] [doi: 10.1093/geront/gny081] [Medline: 29982458]
4. Anderson M, Perrin A. 2. Barriers to adoption and attitudes towards technology. Pew Research Center. 2017. URL: https://www.pewinternet.org/2017/05/17/barriers-to-adoption-and-attitudes-towards-technology/ [accessed 2019-09-14]
5. Berkowsky RW, Sharit J, Czaja SJ. Factors predicting decisions about technology adoption among older adults. Innov Aging 2018;2(1):igy002 [FREE Full text] [doi: 10.1093/geroni/igy002] [Medline: 30480129]
6. Hsieh KL, Fanning JT, Rogers WA, Wood TA, Sosnoff JJ. A fall risk mHealth app for older adults: development and usability study. JMIR Aging 2018;1(2):e11569 [FREE Full text] [doi: 10.2196/11569] [Medline: 31518234]
7. Schumacher RM. Gran got tech: inclusivity and older adults. J Usability Stud 2023;18(2):62-67 [FREE Full text] [doi: 10.1016/b978-0-12-408138-3.00012-1]
8. Sarkar U, Gourley GI, Lyles CR, Tieu L, Clarity C, Newmark L, et al. Usability of commercially available mobile applications for diverse patients. J Gen Intern Med 2016;31(12):1417-1426 [FREE Full text] [doi: 10.1007/s11606-016-3771-6] [Medline: 27418347]
9. Charalambous AP, Pye A, Yeung WK, Leroi I, Neil M, Thodi C, et al. Tools for app- and web-based self-testing of cognitive impairment: systematic search and evaluation. J Med Internet Res 2020;22(1):e14551 [FREE Full text] [doi: 10.2196/14551] [Medline: 31951218]
10. Wald JS, Businger A, Gandhi TK, Grant RW, Poon EG, Schnipper JL, et al. Implementing practice-linked pre-visit electronic journals in primary care: patient and physician use and satisfaction. J Am Med Inform Assoc 2010;17(5):502-506 [FREE Full text] [doi: 10.1136/jamia.2009.001362] [Medline: 20819852]

XSL•FO
RenderX

11. Cusatis R, Holt JM, Williams J, Nukuna S, Asan O, Flynn KE, et al. The impact of patient-generated contextual data on communication in clinical practice: a qualitative assessment of patient and clinician perspectives. Patient Educ Couns 2020;103(4):734-740. [doi: 10.1016/j.pec.2019.10.020] [Medline: 31744702]
12. Gordon NP, Hornbrook MC. Differences in access to and preferences for using patient portals and other eHealth technologies based on race, ethnicity, and age: a database and survey study of seniors in a large health plan. J Med Internet Res 2016;18(3):e50 [FREE Full text] [doi: 10.2196/jmir.5105] [Medline: 26944212]
13. Crouch E, Gordon NP. Prevalence and factors influencing use of internet and electronic health resources by middle-aged and older adults in a US Health Plan population: cross-sectional survey study. JMIR Aging 2019;2(1):e11451 [FREE Full text] [doi: 10.2196/11451] [Medline: 31518256]
14. Irizarry T, Shoemake J, Nilsen ML, Czaja S, Beach S, DeVito Dabbs A. Patient portals as a tool for health care engagement: a mixed-method study of older adults with varying levels of health literacy and prior patient portal use. J Med Internet Res 2017;19(3):e99 [FREE Full text] [doi: 10.2196/jmir.7099] [Medline: 28360022]
15. Harte R, Quinlan LR, Glynn L, Rodríguez-Molinero A, Baker PM, Scharf T, et al. Human-centered design study: enhancing the usability of a mobile phone app in an integrated falls risk detection system for use by older adult users. JMIR Mhealth Uhealth 2017;5(5):e71 [FREE Full text] [doi: 10.2196/mhealth.7046] [Medline: 28559227]
16. Mansson L, Wiklund M, Öhberg F, Danielsson K, Sandlund M. Co-creation with older adults to improve user-experience of a smartphone self-test application to assess balance function. Int J Environ Res Public Health 2020;17(11):3768 [FREE Full text] [doi: 10.3390/ijerph17113768] [Medline: 32466484]
17. Reading Turchioe M, Grossman LV, Baik D, Lee CS, Maurer MS, Goyal P, et al. Older adults can successfully monitor symptoms using an inclusively designed mobile application. J Am Geriatr Soc 2020;68(6):1313-1318 [FREE Full text] [doi: 10.1111/jgs.16403] [Medline: 32157679]
18. Greer DB, Abel WM. Exploring feasibility of mHealth to manage hypertension in rural Black older adults: a convergent parallel mixed method study. Patient Prefer Adherence 2022;16:2135-2148 [FREE Full text] [doi: 10.2147/PPA.S361032] [Medline: 35999840]
19. Feinleib M, Kannel WB, Garrison RJ, McNamara PM, Castelli WP. The Framingham offspring study. design and preliminary data. Prev Med 1975;4(4):518-525. [doi: 10.1016/0091-7435(75)90037-7] [Medline: 1208363]
20. Blair J, Conrad F, Ackermann AC, Claxton G. The effect of sample size on interview findings. AAPOR-American Statistical Association, Section on Survey Research Methods. 2005. URL: http://www.asasrms.org/Proceedings/y2006/Files/JSM2006-000127.pdf [accessed 2024-05-14]
21. Jaspers MWM. A comparison of usability methods for testing interactive health technologies: methodological aspects and empirical evidence. Int J Med Inform 2009;78(5):340-353. [doi: 10.1016/j.ijmedinf.2008.10.002] [Medline: 19046928]
22. Cella D, Choi SW, Condon DM, Schalet B, Hays RD, Rothrock NE, et al. PROMIS adult health profiles: efficient short-form measures of seven health domains. Value Health 2019;22(5):537-544 [FREE Full text] [doi: 10.1016/j.jval.2019.02.004] [Medline: 31104731]
23. Lai JS, Wagner LI, Jacobsen PB, Cella D. Self-reported cognitive concerns and abilities: two sides of one coin? Psychooncology 2014;23(10):1133-1141 [FREE Full text] [doi: 10.1002/pon.3522] [Medline: 24700645]
24. Department of Health Systems and Population Health. Rapid assessment of physical activity (RAPA). University of Washington. 2006. URL: https://depts.washington.edu/hprc/programs-tools/tools-guides/rapa/ [accessed 2024-05-14]
25. Brummett CM, Bakshi RR, Goesling J, Leung D, Moser SE, Zollars JW, et al. Preliminary validation of the Michigan Body Map. Pain 2016;157(6):1205-1212 [FREE Full text] [doi: 10.1097/j.pain.0000000000000506] [Medline: 26835782]
26. Dahmen J, Cook D, Fellows R, Schmitter-Edgecombe M. An analysis of a digital variant of the trail making test using machine learning techniques. Technol Health Care 2017;25(2):251-264 [FREE Full text] [doi: 10.3233/THC-161274] [Medline: 27886019]
27. Ashendorf L, Jefferson AL, O'Connor MK, Chaisson C, Green RC, Stern RA. Trail making test errors in normal aging, mild cognitive impairment, and dementia. Arch Clin Neuropsychol 2008;23(2):129-137 [FREE Full text] [doi: 10.1016/j.acn.2007.11.005] [Medline: 18178372]
28. Hankee LD, Preis SR, Piers RJ, Beiser AS, Devine SA, Liu Y, et al. Population normative data for the CERAD word list and victoria stroop test in younger- and middle-aged adults: cross-sectional analyses from the Framingham Heart Study. Exp Aging Res 2016;42(4):315-328 [FREE Full text] [doi: 10.1080/0361073X.2016.1191838] [Medline: 27410241]
29. Jongstra S, Wijsman LW, Cachucho R, Hoevenaar-Blom MP, Mooijaart SP, Richard E. Cognitive testing in people at increased risk of dementia using a smartphone app: the iVitality proof-of-principle study. JMIR Mhealth Uhealth 2017;5(5):e68 [FREE Full text] [doi: 10.2196/mhealth.6939] [Medline: 28546139]
30. Stoyanov SR, Hides L, Kavanagh DJ, Zelenko O, Tjondronegoro D, Mani M. Mobile app rating scale: a new tool for assessing the quality of health mobile apps. JMIR Mhealth Uhealth 2015;3(1):e27 [FREE Full text] [doi: 10.2196/mhealth.3422] [Medline: 25760773]
31. Fischer B, Peine A, Östlund B. The importance of user involvement: a systematic review of involving older users in technology design. Gerontologist 2020;60(7):e513-e523 [FREE Full text] [doi: 10.1093/geront/gnz163] [Medline: 31773145]

XSL•FO
RenderX

32. Harte R, Hall T, Glynn L, Rodríguez-Molinero A, Scharf T, Quinlan LR, et al. Enhancing home health mobile phone app usability through general smartphone training: usability and learnability case study. JMIR Hum Factors 2018;5(2):e18 [FREE Full text] [doi: 10.2196/humanfactors.7718] [Medline: 29699969]
33. Gomez-Hernandez M, Ferre X, Moral C, Villalba-Mora E. Design guidelines of mobile apps for older adults: systematic review and thematic analysis. JMIR Mhealth Uhealth 2023;11:e43186 [FREE Full text] [doi: 10.2196/43186] [Medline: 37733401]
34. Tech adoption climbs among older adults. Pew Research Center. 2017. URL: http://www.pewinternet.org/2017/05/17/tech-adoption-climbs-among-older-adults/ [accessed 2024-05-14]
35. Ratcliff CL, Krakow M, Greenberg-Worisek A, Hesse BW. Digital health engagement in the US population: insights from the 2018 health information national trends survey. Am J Public Health 2021;111(7):1348-1351. [doi: 10.2105/AJPH.2021.306282] [Medline: 34014759]
36. Scullin MK, Jones WE, Phenis R, Beevers S, Rosen S, Dinh K, et al. Using smartphone technology to improve prospective memory functioning: a randomized controlled trial. J Am Geriatr Soc 2022;70(2):459-469 [FREE Full text] [doi: 10.1111/jgs.17551] [Medline: 34786698]
37. Stockwell S, Schofield P, Fisher A, Firth J, Jackson SE, Stubbs B, et al. Digital behavior change interventions to promote physical activity and/or reduce sedentary behavior in older adults: a systematic review and meta-analysis. Exp Gerontol 2019;120:68-87. [doi: 10.1016/j.exger.2019.02.020] [Medline: 30836130]
38. Riadi I, Kervin L, Dhillon S, Teo K, Churchill R, Card KG, et al. Digital interventions for depression and anxiety in older adults: a systematic review of randomised controlled trials. Lancet Healthy Longev 2022;3(8):e558-e571 [FREE Full text] [doi: 10.1016/S2666-7568(22)00121-0] [Medline: 36102765]
39. Seifert A, Reinwand DA, Schlomann A. Designing and using digital mental health interventions for older adults: being aware of digital inequality. Front Psychiatry 2019;10:568 [FREE Full text] [doi: 10.3389/fpsyt.2019.00568] [Medline: 31447716]
40. Matthew-Maich N, Harris L, Ploeg J, Markle-Reid M, Valaitis R, Ibrahim S, et al. Designing, implementing, and evaluating mobile health technologies for managing chronic conditions in older adults: a scoping review. JMIR Mhealth Uhealth 2016;4(2):e29 [FREE Full text] [doi: 10.2196/mhealth.5127] [Medline: 27282195]
41. Keränen NS, Kangas M, Immonen M, Similä H, Enwald H, Korpelainen R, et al. Use of information and communication technologies among older people with and without frailty: a population-based survey. J Med Internet Res 2017;19(2):e29 [FREE Full text] [doi: 10.2196/jmir.5507] [Medline: 28196791]
42. Zhang S, Boot WR. Predicting older adults' continued computer use after initial adoption. Innov Aging 2023;7(4):igad029 [FREE Full text] [doi: 10.1093/geroni/igad029] [Medline: 37197443]

Abbreviations

eFHS: electronic Framingham Heart Study
FHS: Framingham Heart Study
MARS: Mobile App Rating Scale
PROMIS: Patient Reported Outcomes Measurement Information System
REDCap: Research Electronic Data Capture

Edited by A Choudhury; submitted 22.01.24; peer-reviewed by D Chong, C Moral; comments to author 29.03.24; revised version received 28.04.24; accepted 01.05.24; published 30.05.24.

Please cite as:
Murabito JM, Faro JM, Zhang Y, DeMalia A, Hamel A, Agyapong N, Liu H, Schramm E, McManus DD, Borrelli B
Smartphone App Designed to Collect Health Information in Older Adults: Usability Study
JMIR Hum Factors 2024;11:e56653
URL: https://humanfactors.jmir.org/2024/1/e56653
doi:10.2196/56653
PMID:38815261

©Joanne M Murabito, Jamie M Faro, Yuankai Zhang, Angelo DeMalia, Alexander Hamel, Nakesha Agyapong, Hongshan Liu, Eric Schramm, David D McManus, Belinda Borrelli. Originally published in JMIR Human Factors (https://humanfactors.jmir.org), 30.05.2024. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information,

XSL•FO
RenderX

a link to the original publication on https://humanfactors.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

The Influence of Incentive-Based Mobile Fitness Apps on Users' Continuance Intention With Gender Moderation Effects: Quantitative and Qualitative Study

Aaya Faizah[1], SKom; Alifah Fatimah Azzahra Hardian[1], SKom; Rania Devina Nandini[1], SKom; Putu Wuri Handayani[1], Prof Dr; Nabila Cyldea Harahap[1], Dr

Faculty of Computer Science, University of Indonesia, Depok, Indonesia

Corresponding Author:
Putu Wuri Handayani, Prof Dr
Faculty of Computer Science, University of Indonesia
Kampus UI
Depok, 16424
Indonesia
Phone: 62 217863419
Email: Putu.wuri@cs.ui.ac.id

Abstract

Background: A survey conducted by McKinsey & Company reported that, as of May 2022, as many as 26% of Indonesians had recently started to engage actively in physical activity, 32% undertook regular physical activity, and 9% exercised intensely. The Fourth Industrial Revolution has spurred the rapid development of mobile fitness apps (MFAs) used to track people's sports activities. However, public interest in using these apps for any length of time is still relatively low.

Objective: In this study, we aimed to determine the effect of incentives (eg, self-monitoring, social support, platform rewards, and external influence) on the use of MFAs and the moderating effect of gender on users' continuance usage intention.

Methods: The study used a mixed methods approach. Quantitative data were collected through a web-based questionnaire and qualitative data from interviews with 30 respondents. The quantitative data, collected from 379 valid responses, were processed using covariance-based structural equation modeling. The qualitative data were processed using thematic analysis. The MFAs included in this research were those used as sports or physical activity trackers, such as Apple Fitness, Strava, Nike Run Club, and Fita.

Results: The results of the data analysis show that 3 groups of incentives, namely, self-monitoring, platform rewards, and external influence (with the exception of social support), affect the perceived usefulness of these apps. Gender was also shown to moderate user behavior in relation to physical activity. The study showed that women were more likely to be motivated to exercise by social and external factors, while men paid greater attention to the tracking features of the app and to challenges and rewards.

Conclusions: This research contributes to the field of health promotion by providing guidance for MFA developers.

(JMIR Hum Factors 2024;11:e50957) doi:10.2196/50957

KEYWORDS

incentive; fitness; mobile fitness apps; gender; continuance usage intention; Indonesia; mobile phone

Introduction

Background

According to the World Health Organization [1], regular physical activity (PA) is a key factor in the prevention and management of noncommunicable diseases. The Global Status Report on Physical Activity [2] reported that 1.4 billion individuals aged >18 years do not meet the levels of PA recommended to promote and protect health. In 2016, it was reported that, globally, 23% of all men and 32% of all women aged ≥18 years were not sufficiently physically active to stay healthy [1]. This means that approximately 1 in 3 women and 1 in 4 men are not sufficiently active and do not meet the global recommendation of at least 150 minutes of moderate-intensity PA or 75 minutes of high-intensity PA per week [3]. In August 2022, McKinsey & Company released the results of a survey conducted with 1041 Indonesian respondents in which 26% of

XSL•FO
RenderX

the respondents stated that they had started to engage actively in personal training, 32% reported that they had been playing sports regularly, and 9% indicated that they had increased the intensity of their sports or fitness activities [4]. These data indicate that the level of Indonesians' interest in, and awareness of, sports, fitness, and personal training is significantly higher than the global average. This conclusion is supported by the increased use of mobile fitness apps (MFAs) in Indonesia, with 29 million users in 2022 [5].

MFAs use several types of incentives, which include self-incentives, peer incentives, and platform incentives [6-9]. In the MFA context, self-incentives involve a self-monitoring (SM) system in which users monitor and track their own behavior [10,11]. Peer incentives are focused on social support (SS), which includes informational, emotional, and material support or the protection provided by fellow users of the app [12]. In the context of MFAs, platform incentives usually take the form of rewards or awards resulting from gamification features [13]. Users who collect a large number of rewards are usually considered to have a higher status on the MFA and feel more satisfied with their use of the app [14,15].

According to Zhu et al [16], very few studies have examined the role played by gender differences in the use of health and fitness apps. Yin et al [17] stated that achievements in sports motivate men more, while social relationships motivate women more. Previous research on MFAs has explored their design [18] and evaluation [19-22], as well as user adoption intentions [23]. In addition, several studies have discussed continuity in the use of MFAs [24-27]. Chiu et al [25] integrated the expectation-confirmation theory (ECT) with the investment model to analyze the continuous use of MFAs. However, research investigating the various types of MFA incentives has been shown to have several limitations [17,26] because the effects of each incentive have mostly been explored separately [28-30]. Per McKinsey & Company's 2022 survey among Indonesian citizens [4], 87% of the respondents intended to continue using their personal training and fitness apps. The market analysis and demographics of this study apply only to MFA users in Indonesia.

Research Question

This research adopted the self-determination theory (SDT) and the ECT. The SDT, as postulated by Ryan and Deci [31], states that there are 3 main psychological needs that drive human behavior: autonomy, relatedness, and competency. If these psychological needs are met, intrinsic motivation will increase and make it easier to maintain certain behaviors [31]. Teixeira et al [32] show that the SDT can be applied to behavioral interventions that relate to exercise or PA. While the SDT has the ability to predict the intensity of a behavior based on the influence of incentive factors [17], the ECT is generally used to predict the continuity of a behavior [25]. The combination of the SDT and the ECT was chosen to analyze the relationship between the incentive factors that affect the use of MFAs and continuity in using them. Thus, the research question is "How do the incentives promoted by MFAs influence users' continuance usage intention (CUI)?" This research can provide guidance for MFA developers by helping them to evaluate their apps.

Methods

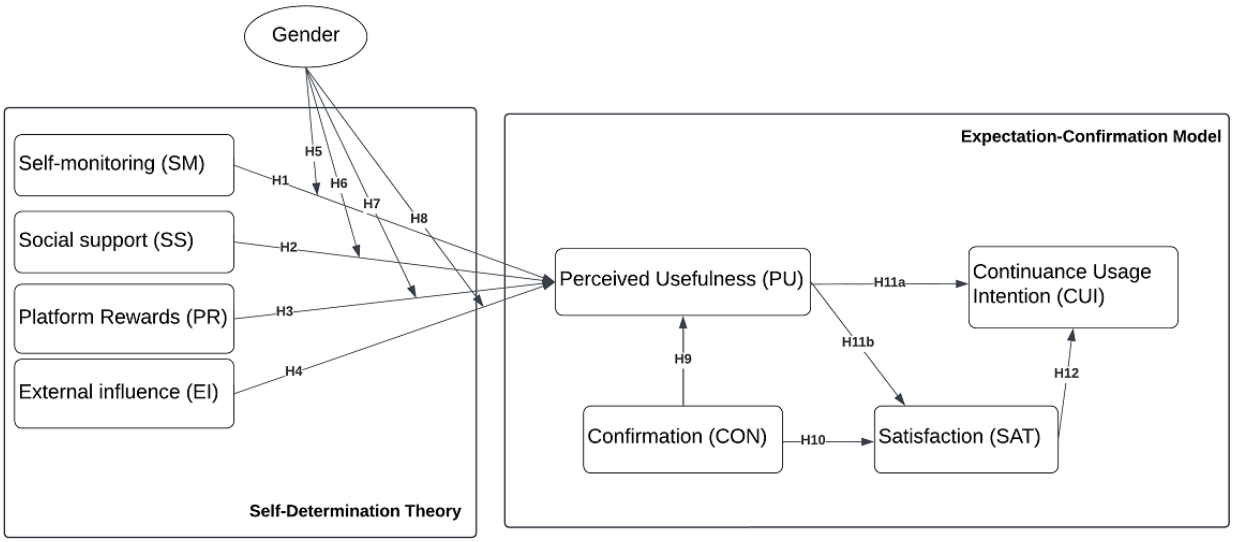
Research Model

Overview

The model used for this research is based on 2 theories and 1 moderating effect, namely, the SDT and the ECT, with the moderating effect of gender. Significant studies reporting on the use of these 2 theories include those by Yin et al [17], Huang and Ren [26], Chiu et al [25], and Li et al [33]. Yin et al [17] found that incentives are compatible with the SDT in motivating users' PA behaviors. The SDT approach described by Yin et al [17] is the theoretical basis for this research because it analyzes incentives offered by MFAs collectively and uses gender as a moderating variable. The relationship between perceived usefulness (PU) and incentives was also analyzed by Huang and Ren [26]. This research suggests that technology functions in MFAs, such as SM, self-regulation, and goal attainment, have an indirect effect on CUI through PU; for instance, Chiu et al [25] and Li et al [33] found that users' CUI was significantly predicted by ECT. Our research model, which includes 9 variables and 13 hypotheses (described in the following subsections), is presented in Figure 1.

XSL•FO
RenderX

Figure 1. Proposed conceptual model.



The Influence of SM on PU

SM, which is classified as one of the self-incentives in MFAs, includes managing and tracking one's own behavior [17]. These actions enable users to observe their own progress and evaluate their performance against previously set goals [34]. PU refers to the extent to which a person feels that technology can improve their performance of certain tasks [35]. In this study, the task was identified as increasing the user's PA, while, for MFA users, PU implies that using the MFA will enhance their personal training intensity [36,37]. Bhattacherjee [38] argues that when users confirm their initial expectations of the main functionality of a mobile app, they will begin to perceive the app as useful for improving their task performance and thus continue to use it. Huang and Ren [26] measured PU relating to the effectiveness and performance of PA through the use of 4 technological functions of the MFA, one of which is SM. Therefore, we examine the following hypothesis:

- H1: SM has an influence on PU.

The Influence of SS on PU

SS is classified as one of the peer incentives in MFAs [17]. Web-based SS is seen as an important factor affecting the physical and mental health of individuals, such as sports activity and increased well-being [12,39]. Humans have a tendency to behave in ways that are consistent with people in their own social networks, and this can be exploited in the context of mobile health (mHealth) [29]. Chen and Pu [40] conducted research on social incentives by developing the HealthyTogether mobile game, which allows users to participate in PA together and send messages to one another. The authors showed that users significantly increased their PA when using HealthyTogether compared to when they were exercising alone [40]. Edney et al [6] built the Active Team app, which is an MFA with social and gamification functions. The primary outcome of their study was a change in the total daily minutes of moderate to vigorous PA at 3 months, as measured objectively using an accelerometer [6]. Therefore, we propose the following hypothesis:

- H2: SS has an influence on PU.

The Influence of Platform Rewards on PU

The platform rewards include gamification elements, such as badges, points, and leaderboards [17]. The gamification element in MFAs can provide two types of information: (1) the user's PA progress and (2) a comparison of the user's PA with that of other users [41]. From this information, MFA users can observe their progress and experience greater satisfaction as they recognize their own personal training achievements. This leads to higher user competency satisfaction and increased behavioral motivation [13,42]. Yin et al [17] found that platform rewards have a positive relationship with users' PA. This finding is supported by Plangger et al [13] and Huang and Ren [26], who analyzed the effect of the goal-attainment technology function of MFAs, in which users can set their own goals, which are then achieved by undertaking PA. These achievements are then categorized as platform rewards. Huang and Ren [26] also found that this technology function had a positive effect on PU. Therefore, we propose the following hypothesis:

- H3: Platform rewards have a positive influence on PU.

The Influence of External Influence on PU

External influence (EI) is one of the extrinsic motivations identified in the SDT, which means that behavior is motivated through influences that do not depend on internal factors [43]. Huang [28] proposed this variable to explain how PA can be promoted through external factors. One example is companies providing incentives to MFA users as part of their corporate social responsibility initiatives [28]. Several studies discuss EI and PA. One example of EI referred to in this study is the name or *image* of a sponsor of an activity [44]. Low and Pyun [45] explain that sponsorship that gives a good impression to customers or users will produce behavior that tends to be positive. In the context of sporting activities, Huang [28] explains that sponsor characteristics play an important role in participation in a sporting activity. Therefore, because an MFA is a tool that can measure a person's PA, we intend to explore the following hypothesis:

XSL•FO
RenderX

- H4: EI has an influence on PU.

The Influence of Gender on SM and PU

According to Mao et al [7], MFA incentives are not always equally effective for women and men. This is because women and men have different ways of thinking [17]. Yin et al [17] conducted research that assumed that gender would influence the effectiveness of SM incentives, making them more effective for men than for women. This assumption was based on the belief that men generally pay more attention to their own achievements than women [46]. Surprisingly, Yin et al [17] show that gender does not affect the effectiveness of SM in MFAs. This finding relates to the concept of self-regulation, which is strongly driven by self-efficacy [47]. Individuals who decide to use MFAs are generally believed to have high self-efficacy in carrying out PA [17]. Therefore, we plan to test the following hypothesis:

- H5: Gender influences the relationship between SM and PU in MFA users.

The Influence of Gender on SS and PU

With regard to SS, Yin et al [17] explain that SS is one of the factors that most helps to fulfill the relatedness needs described in the SDT. According to Wang et al [9], social ties and commitment are more important for women than for men in shaping their attitudes toward the sharing of information. In considering gender, Yin et al [17] found that women tend to be more influenced by their relatedness needs than men. Women are also believed to be driven more by collective goals, such as pleasure or interpersonal harmony [48,49]. In the context of health apps, Kimbrough et al [50] found that women are usually more affected by environmental conditions and social relationships than men. Thus, we propose the following hypothesis:

- H6: Gender influences the relationship between SS and PU among MFA users.

The Influence of Gender on Platform Rewards and PU

Men tend to focus more on themselves and tend to be more independent than women [46,51]. Men also tend to focus more on completing or achieving individual goals that demonstrate their performance and abilities [46,51]. Relatedly, Vilela and Nelson [52] showed that men tend to be more motivated by their own achievements than women when using information system products. This is due to the general behavioral characteristics of men, who are generally more aggressive, pragmatic, and self-oriented in their behavior compared to women [52]. When specifically applied to incentives and CUI, Yin et al [17] also found an influence between gender and the effectiveness of platform reward incentives. The authors assumed that this is caused by the behavioral characteristics of men, who generally make decisions more rationally and pay greater attention to their own behavior. Thus, we propose the following hypothesis:

- H7: Gender influences the relationship between platform rewards and PU among MFA users.

The Influence of Gender on EI and PU

Sun and Zhang [53] state that women have a higher awareness of the environment than men. Leong et al [54] and Li et al [33] also found that men tend to be less easily influenced by external advice or support. Similarly, Venkatesh et al [55] concluded that women tend to be more influenced by EI, while men are usually less affected by external facilitation in their use of technology. This was confirmed by Weman Josefsson et al [56], who showed that men participate in challenges organized by the community to compete, while women participate for social and autonomy reasons. Hence, we propose the following hypothesis:

- H8: Gender influences the relationship between EI and PU among MFA users.

The Influence of Confirmation of Expectations on PU

Confirmation of expectations refers to the perceived level of conformity between the information system product or service expectation and actual performance [38]. Bhattacherjee [38] explains that PU refers to the individual's perception of the anticipated benefits from the use of IT products or services. The ECT implies that the confirmation of a user's expectations has a positive effect on their perception of the PU of an IT product or service [25,57-59]. According to the cognitive dissonance theory [60], IT users may experience psychological conflict if their initial expectations are not confirmed by their actual use experience [61]. Conversely, if users' initial expectations are confirmed or met, they may display higher investment behavior and reduce their preference for alternative apps [25]. Hsu and Lin [62] state that confirmation of expectations is positively related to the perceived quality of the IT product or service used, with the result that users tend to ignore quality alternatives. Therefore, we propose the following hypothesis:

- H9: Confirmation of expectations has an influence on PU.

The Influence of Confirmation of Expectations on Satisfaction

Chiu et al [25] proposed that confirmation of user expectations affects satisfaction with the app as well as its PU. Satisfaction can be interpreted as an individual's evaluation of their initial experience with a product or service [38]. Chiu et al [25] explain that before downloading an app, users generally have expectations of it, based on detailed information received from the app provider and on ratings and reviews from other users. After using the app, the user gains experience and evaluates the performance of the app based on previously established expectations. In line with the expectation-confirmation model, Chiu et al [25] assume that users' perceptions of postuse benefits and the confirmation of previous expectations determine their satisfaction in using IT products and services. Therefore, we propose the following hypothesis:

- H10: Confirmation of expectations has an influence on satisfaction.

The Influence of PU on CUI

PU refers to the user's perception of the benefits expected from using an IT product or service [61,63]. According to Bhattacherjee [38], expectations based on the user's direct

XSL•FO
RenderX

experience have an important role in forming their IT CUI. Chiu et al [25] state that many studies conducted in various contexts [59,64,65] empirically support a positive relationship between PU and CUI. Wu et al [66] show that when users find the mHealth app useful, they show a higher level of satisfaction and tend to use it continuously. Thus, we define the following hypothesis:

- H11a: PU has an influence on CUI.

The Influence of PU on Satisfaction

According to Chiu et al [25], PU also has a strong and positive impact on satisfaction. The authors state that the more benefits users receive from health and fitness apps, the greater their satisfaction [25]. When a user has used an app for an extended period of time, the user will evaluate its performance and form either a confirmation or a disconfirmation of judgment with regard to their expectations [62]. Disconfirmation of expectations affects user satisfaction and creates negative perceptions of the usefulness of MFAs. Conversely, users' positive perceptions of usefulness increase their satisfaction with an app. Therefore, we propose the following hypothesis:

- H11b: PU has an influence on satisfaction.

The Influence of Satisfaction on CUI

Satisfaction can be identified as a significant factor influencing consumer behavior [25]. Bhattacherjee [61] strengthens this definition by explaining that user satisfaction is an important determinant of postadoption behavior relating to IT products and services. In other words, users with higher levels of satisfaction will exhibit greater levels of use of IT products and services than those who are less satisfied [25]. Wu et al [66] confirm that satisfied users are more likely to continue using an app because dissatisfied users can easily switch to other technologies at no additional cost. The relationship between satisfaction and CUI has been identified as one of the strongest relationships in the expectation-confirmation model [63]. Therefore, we propose the following hypothesis:

- H12: Satisfaction has an influence on the CUI of MFA users.

Research Procedure

This study used a mixed methods approach that integrated a quantitative approach, based on a questionnaire, with a qualitative approach, using interviews. The only inclusion criterion for respondents in this study was that they used MFAs. We modified a questionnaire that has been established in previous studies [12,15,16,24,25,28,33,66-70]. Before distributing the questionnaire, a readability test was conducted to validate how easily the questionnaire could be understood by respondents. The readability test was carried out both face-to-face and internet-based, using Google Meet, with 8 people who met the research criterion (ie, they all used MFAs). This readability test was carried out between February 5 and 10, 2023. We then used the results of the readability test to refine the questionnaire.

Once the questionnaire had been refined, we conducted a pilot study from February 20 to 25, 2023, aiming to measure the validity and reliability of the questionnaire by distributing it to 31 selected research respondents. The results of the pilot study were used to check the value of Cronbach α, which, in this pilot study, was 0.832, well over the required value of >0.7.

Research Instruments

The instruments used in this study were a web-based questionnaire and semistructured interview questions. The questionnaire first asked questions regarding the demographics of the respondents, and it then presented statements regarding the research model being tested. Each of the 8 variables exclude the gender variable in the study was assessed by 3 or 4 measurement items, and each indicator was represented by a statement to which participants responded on a Likert scale ranging from 1=*strongly disagree* to 5=*strongly agree*. The questionnaire used in this study is available in Multimedia Appendix 1, and a list of the interview questions is available in Multimedia Appendix 1.

Ethical Considerations

This research was approved by Faculty of Computer Science (approval number S-7/UN2.F11.D1.5/PPM.00.00/2024).

Results

Participant Demographics

We distributed the research questionnaire on the web through various social media platforms such as WhatsApp, Line, Twitter, Instagram, and Telegram. These social media platforms are widely used by Indonesians. The questionnaire distribution was carried out between February 27 and March 20, 2023. Table 1 provides a demographic summary of the respondents. Of the respondents, 75.5% (286/379) were aged between 17 and 25 years, 72.3% (274/379) were women, 25.1% (95/379) were privately employed, and 51.5% (195/379) lived in Greater Jakarta.

XSL•FO
RenderX

Table 1. Respondents' demographics (n=379).

Variable	Respondents, n (%)
Age (years)	
17-25	286 (75.5)
26-35	74 (19.5)
36-45	14 (3.7)
>45	5 (1.3)
Gender	
Woman	274 (72.3)
Man	105 (27.8)
Occupation	
College student	201 (53)
Employee of state-owned enterprise	8 (2.1)
High school student	15 (4)
Privately employed	95 (25.1)
Unemployed	19 (5)
Entrepreneur	23 (6.1)
Housewife	2 (0.5)
Civil servant	3 (0.8)
Other	13 (3.4)
Domicile	
Greater Jakarta	195 (51.5)
Java island	137 (36.1)
Outside of Java island	47 (12.4)

After collecting both the quantitative and qualitative data, we processed the quantitative data using covariance-based structural equation modeling. Using covariance-based structural equation modeling, data processing is carried out in several stages: specification and identification of the research model, estimation of the research model, testing the feasibility of the research model, modification of the research model, and hypothesis testing.

To validate the quantitative data results, we also collected qualitative data by conducting semistructured interviews with 30 respondents. The interviews were conducted both offline and on the web and took 30 to 45 minutes each. The qualitative data analysis was carried out thematically on the basis of the defined hypotheses.

Measurement Model

The factor loading values of all variables and indicators met the Cronbach α standard of >0.7 [71]; thus, the model feasibility test could be carried out. This study yielded average variance extracted values >0.5 as well as Cronbach α and composite reliability values >0.7 [71] (Table 2).

Table 2. Average variance extracted, Cronbach α, and composite reliability values.

Variable	Average variance extracted	Cronbach α	Composite reliability
Self-monitoring	0.968	0.705	0.920
Platform rewards	0.773	0.927	0.872
External influence	0.865	0.816	0.834
Social support	0.638	0.854	0.835
Confirmation of expectations	0.975	0.763	0.885
Perceived usefulness	0.709	0.888	0.848
Satisfaction	0.669	0.889	0.889
Continuance use intention	0.640	0.812	0.842

XSL•FO
RenderX

Structural Model

Next, we tested the structural model with the goodness-of-fit criteria, which included the relative chi-square index, goodness-of-fit index, root-mean-square error of approximation, root mean square residual, normal fit index, comparative fit index, and the Tucker-Lewis Index [71]. The goodness-of-fit values are presented in Table 3, and the R^2 values are shown in Table 4.

Table 3. Goodness-of-fit values.

Goodness-of-fit criteria	Cutoff value	Value	Description
Relative chi-square index	<2	1.956	Good fit
Goodness-of-fit index	≥0.9	0.900	Good fit
Root-mean-square residual	≤0.05	0.048	Good fit
Normal fit index	≥0.9	0.913	Good fit
Comparative fit index	≥0.9	0.955	Good fit
Tucker-Lewis Index	≥0.9	0.948	Good fit
Root-mean-square error of approximation	≤0.08	0.050	Good fit

Table 4. R^2 values.

Variable	R^2	Effect size
Perceived usefulness	0.349	Weak
Satisfaction	0.511	Medium
Continuance use intention	0.714	Strong

Hypotheses Testing

This study used a 2-tailed significance test; thus, the condition for accepting the hypothesis was $P<.05$ [71]. Table 5 presents the results of hypotheses 1 to 4 and 9 to 12, only one of which (H2) was rejected.

Table 5. Hypotheses testing results.

Hypothesis	Estimate (95% CI)	*P* value	Result
H1: SM[a]→PU[b]	0.319 (0.244 to 0.394)	.001	Accepted
H2: SS[c]→PU	0.060 (−0.019 to 0.143)	.12	Rejected
H3: PR[d]→PU	0.136 (0.044 to 0.219)	.007	Accepted
H4: EI[e]→PU	−0.101 (−0.166 to −0.033)	.006	Accepted
H9: COE[f]→PU	0.323 (0.251 to 0.388)	.001	Accepted
H10: COE→satisfaction	0.541 (0.435 to 0.632)	.002	Accepted
H11a: PU→CUI[g]	0.280 (0.200 to 0.363)	.002	Accepted
H11b: PU→satisfaction	0.218 (0.069 to 0.355)	.003	Accepted
H12: Satisfaction→CUI	0.683 (0.560 to 0.813)	.001	Accepted

[a]SM: self-monitoring.
[b]PU: perceived usefulness.
[c]SS: social support.
[d]PR: platform rewards.
[e]EI: external influence.
[f]COE: confirmation of expectations.
[g]CUI: continuance usage intention.

According to Awang [72], the test for moderation is not significant when the difference in chi-square values between the constrained model and the unconstrained model is <3.84. Table 6 presents a summary of the results of the hypothesis testing using the moderating effect of gender. On the basis of the difference in the chi-square values between the constrained model and the unconstrained model, it can be concluded that all difference values were >3.84 and therefore meet the

XSL•FO
RenderX

requirements for calculating the significance of the moderating effect, meaning that H5, H6, H7, and H8 were all accepted.

Table 6. Summary of moderating variable hypothesis testing, with gender as the moderating effect.

Path	Chi-square constrained model	Chi-square unconstrained model	Difference (*df*)	Result
SM[a]→PU[b]	1318.671	1043.151	275.520	H5 accepted
SS[c]→PU	1627.823	1043.151	584.672	H6 accepted
PR[d]→PU	1395.076	1043.151	351.925	H7 accepted
EI[e]→PU	1796.875	1043.151	753.724	H8 accepted

[a]SM: self-monitoring.
[b]PU: perceived usefulness.
[c]SS: social support.
[d]PR: platform rewards.
[e]EI: external influence.

Qualitative Interviews and Validity of the Hypotheses

This research shows that the incentives offered in MFAs in the form of SM (eg, distance walked or run, number of calories expended, time taken, and heart rate) influence users' motivation to undertake PA. The acceptance of H1 is thus in accordance with the findings of Yin et al [17] and Stragier et al [73]. Yin et al [17] state that the user's PA level correlates positively with the amount of SM they do. The majority of interviewees felt that the SM feature provided encouragement for their PA:

> *So I feel happy because I have exercised, more enthusiasm.* [Interviewee 6]

In addition, the interviewees believed that MFAs documented or tracked their progress in PA, which helped them to maintain or even improve their exercise consistency:

> *So that I can compare with previous progress and so that in the future I can look back at my history. Like pace, I also remember what date I did sport.* [Interviewee 9]

An example of a feature that can be implemented is one that displays a summary of the user's performance while exercising, together with visualizations in the form of trends and graphs. Some apps also display comments that describe the user's sports activity performance, based on their activity level. Users can take advantage of these insights to increase their PA levels in their next sports activity.

However, H2 was rejected in this study. H2's rejection aligns with the findings of Sun and Jiang [74] and Kim et al [75]. According to Kim et al [75], social comparison and the user's level of PA are not directly connected. Social comparison here is defined as the relationship between the level of PA and the variable self-efficacy, or a person's belief in their own capabilities [75]. The rejection of H2 indicates that the community or social ecosystem around MFA users does not have a significant impact on motivating the users to exercise. On the basis of the interviews, the SS feature in the app does not have an important effect on PA levels because users do not feel compelled to exercise when using the SS feature:

> *There is no motivation from the engagement side, more from tracking my own progress.* [Interviewee 9]

In addition, nearly one-third of the interviewees (9/30, 30%) admitted that they used the SS feature only to document sports activities that had already been completed.

H3 was accepted in this study. The acceptance of H3 aligns with the studies by Bojd et al [41], Payne et al [42], Plangger et al [13], Goes et al [76], and Hamari and Koivisto [77]. Bojd et al [41] found that the gamification element in MFAs can provide two types of information: (1) the user's PA progress and (2) a comparison of the user's PA with that of other users. Furthermore, when MFA users are able to observe their progress, they feel more satisfied and recognize their own PA competency, which will drive higher user competency satisfaction and behavioral motivation [13,42]. Goes et al [76] and Hamari and Koivisto [77] also highlight the gamification element in MFA, which tracks the user's effort, progress, and achievement of personal goals. According to Goes et al [76], the public nature of user-acquired gamification elements, such as levels, badges, or leaderboards, can generate users' social status on the MFA platform, which encourages social comparison and competitive motivation among users. On the basis of the interviews, MFA users want to take part in challenges (an example of implementing gamification) on the app because they want to obtain limited edition rewards and measure their own capabilities in sports activities:

> *Gamification keeps me motivated and helps me see my activities historically during physical activity based on the badge I have earned.* [Interviewee 21]

Furthermore, the interviewees acknowledged that the rewards they obtain can be used as a benchmark of their capacity in the sports activity against which to build new achievements:

> *I feel happy when I get an achievement because it shows an improvement in my sport. Even though I don't I have specifically targeted certain achievements, but if I can surpass the previous achievements, it means that my sport has improved. The goals that I have set are higher than before.* [Interviewee 17]

XSL•FO
RenderX

Therefore, it would be better if the MFA included challenges that were personalized as well as recommendations that were based on the user's type of sports activity, the user's sports activity goals, and the user's own sports activity history. An example of such a feature could be that, based on the user's history, if they have only managed to run a distance of 3 km, then, to improve their performance, other MFA users could recommend a 4-km challenge.

Huang [28] found that sponsor characteristics play an important role in triggering user behavior. Sponsorship referred to circumstances where the use of a sponsor's product occurred naturally as part of a sponsored event [78]; for example, with an MFA whose function is to promote PA, sponsorship of athletic apparel would be perceived as highly congruent, whereas sponsorship of a cold remedy would reflect low congruence. The H4 finding is in line with the study by Yang et al [79], who stated that the level of involvement of a brand produces a positive association with the brand and strengthens the positive effect of an evaluation impacting one's behavioral intention toward an app. From the interviews, it was found that interviewees were encouraged to take part in a challenge or activity if the activity was associated with the party (public figure, company, etc) that organized it:

> *For a club other than Strava, I think it's cool if you participate, for example, it's like unique. There's definitely a challenge made by Strava every month, so it's not as special as other clubs. The limited edition is more about Heart Month, New Year, and others. I want to take part because it would be a shame if I didn't follow.* [Interviewee 9]

We found that not many interviewees took advantage of EI incentives, but those who did participate focused more on the challenges than on the organizers or the external community. If a user felt capable of taking part in a challenge, they would try to do so:

> *Actually, I see from the challenge, if I feel capable, then I want to join.* [Interviewee 7]

Thus, we argue that it would be better if MFA developers or providers developed challenges for their apps that are created by communities, organizations, and figures with high functional congruence.

With regard to H5, H6, H7, and H8, the results show that, in every case, gender has a moderating effect on the relationship between the variables investigated. This study showed that gender influences the relationship between SM and PU (H5). These results are supported by the studies by Gabriel and Gardner [46] and Sun et al [51], who found that men tend to make decisions based on rationality, while women tend to be more perceptual. According to Gabriel and Gardner [46] and Sun et al [51], men are generally more focused on personal goals that demonstrate their individual performance and abilities, while women are usually less conscious of their own goals and performance. This finding is supported by van Elburg et al [80], who state that men focus more on practical goals and achieving goals when using an mHealth app. We found that our female interviewees usually used the metrics in MFAs for tracking their PA only as monitoring information:

> *I only look at the pulse.* [Interviewee 1]

However, the men usually used these metrics as targets for self-development:

> *To find out whether in sports we have reached the desired target or not. On the other hand, if our sports performance is good, this can also be seen through the information displayed on Apple Watch. Thus, the Apple Watch can be a helpful tool in determining whether our performance has reached the expected level or not.* [Interviewee 12]

Moreover, this study found that gender influenced the extent to which SS incentives affected users' PU (H6). The results of the interviews showed that most female respondents felt more motivated by their social community or by the SS feature provided in the MFA they used. By contrast, the male users used the SS feature, such as sharing their sports activity progress, for personal documentation purposes:

> *Just so you know. Only for review, not to share with other friends.* [Interviewee 28]

Other male respondents stated that this was the case simply because the app posted their activity automatically*:*

> *Because it has to be posted on the Strava application.* [Interviewee 8]

Many male respondents had never used this feature, indicating their lack of interest in the SS feature:

> *I have never tried it.* [Interviewee 18]

However, the female respondents all expressed interest in the SS feature available in MFAs and felt more motivated to exercise due to this feature:

> *I also become motivated to exercise when I see my friends after posting their sports results.* [Interviewee 11]

Some of the female respondents commented that the SS feature of MFAs motivated them to exercise by creating a sense of competition:

> *If I just wake up in the morning and get a notification that my friend has finished exercising, I feel left behind because I just woke up but he has finished exercising. Section it motivates, really.* [Interviewee 9]

Relatedly, Li et al [33] found that women pay greater attention to social relations and are more willing to accept support from those around them. By contrast, Leong et al [54] found that men usually ignore external advice or support due to their sense of independence. These findings are supported by Yin et al [17], who found that SS had a more positive effect on PA in women than in men.

This study also showed that gender influences the relationship between platform rewards and PU (H7). The interviews showed that male respondents were generally more motivated by the challenges, badges, and awards offered by the MFA they were using:

XSL•FO
RenderX

> *Makes me more enthusiastic for the next run, and I use it to keep track of whether I should improve or maintain, for example, I can rank third so I feel I have to improve my performance.* [Interviewee 16]

However, the female respondents usually followed or used this feature only for their own satisfaction and without specific targets or motivations:

> *There is no specific goal to get rewards, but I feel happy and proud of myself if I get them.* [Interviewee 13]

According to Yin et al [17], in the context of PA, men usually pay more attention to meeting their needs for autonomy and competence, such as badges, awards, and so on. This was also demonstrated by Vilela and Nelson [52], who stated that men tend to be more aggressive, pragmatic, and self-oriented. Therefore, they are motivated by the need for achievement when using information system products [52]. Similar findings were identified by Forman et al [81], who showed that the gamification element has a more positive effect on men than on women by arousing their competitive and achievement-oriented motivation. Brandts et al [82] also support this finding and explain that task-based goal setting increases task completion and performance only for men.

This study also showed that gender affected the impact of EI on the user's PU of MFAs (H8). The results of the interviews confirmed that there are 2 main reasons a person will participate in PA supported by the MFA: the match between the organizer of the activity and the user and the match between the user's capabilities and the activity or challenge created. Comparing these 2 reasons, we found that the women were more likely to do something because of a match with the organizers, in contrast to the men, who usually focused more on their own ability to participate in an activity:

> *If Strava doesn't have motivation, if it's a club other than in my opinion, Strava is cool if you join, it's like unique. What Strava makes is there every month, so it's not as special as other clubs to participate on Strava.* [Interviewee 29]

According to Huang [28] and Yang et al [79], the reason female MFA users participate in sports activities is that they experience a *special* feeling because these sports activities are created by a special club. Huang [28] and Yang et al [79] explain that the sponsorship characteristics of a sports activity and high brand involvement play important roles in triggering the behavioral intention of MFA users and their behavior in general. H8 is also supported by the findings of Weman Josefsson et al [56], who explain that men tend to be more influenced by winning rewards than women, who tend to participate more for autonomous and social reasons.

Furthermore, H9 was confirmed in this study. The acceptance of H9 is in accordance with previous research conducted by Bhattacherjee [38], Huang et al [15], Chiu et al [25], Wang et al [9], Cai et al [83], and Wu et al [66]. Wu et al [66] found that PU and user satisfaction are directly influenced by confirmation of expectations, namely, the realization of the expected benefits of using mHealth. This result is supported by Chiu et al [25], who state that PU of the MFA is reflected in the user's enhanced exercise capacity and satisfaction, as evidenced by their increased enjoyment of exercising. Thus, it is to be expected that, after the initial experience, the confirmation level of the user's expectations will have a positive effect on their PU [9,15,38,83]. One of the expectations of a respondent who used an MFA was that they would experience changes and improvements in their PA or exercise, and these expectations were indeed successfully confirmed:

> *Because when I want to download Strava I want to be diligent in exercising, and it is proven that I exercise more often because I can track my sports progress.* [Interviewee 23]

H10 was also accepted by this study, and this result is in accordance with the studies by Bhattacherjee [38], Huang et al [15], Chiu et al [25], Wang et al [9], Cai et al [83], and Wu et al [66]. Wang et al [9] found that confirmation of expectations positively affects user satisfaction with IT products and services. The results of the interviews confirmed that interviewees felt satisfaction when using MFAs:

> *From a user point of view, everything has been fulfilled in my opinion. What I need so far has been achieved.* [Interviewee 24]

> *In my opinion, the features are quite complete, because that's all I really need. The application also provides a reminder if you have passed one day without exercising and automatically arranges for the workout that can be fulfilled the next day to be even tougher.* [Interviewee 10]

This study also showed that PU influences CUI. Acceptance of H11a is in accordance with the studies by Bhattacherjee [38], Huang et al [15], Chiu et al [25], Huang and Ren [26], Wang et al [9], Cai et al [83], Wu et al [66], and Cho et al [24]. Cho et al [24] reported that, in the context of MFAs, perceived benefits were associated with managing health-related information. The interviews confirmed that interviewees would continue using the MFAs if they helped them to be more active in their exercising, and they could track their sports activity progress effectively:

> *I will continue to use it because in my opinion it is also effective and looks simple.* [Interviewee 10]

> *As long as device is connected to the Apple Watch, will still use it. The ability to track different types of exercise separately is one of the advantages of the Apple Watch. This makes me still choose to use the Apple Watch in the future, as long as it meets my sporting needs.* [Interviewee 12]

The study's acceptance of H11b is in accordance with the studies by Bhattacherjee [38], Huang et al [15], Chiu et al [25], Wang et al [9], Cai et al [83], and Wu et al [66]. Cai et al [83] explain that PU is reflected in user satisfaction when exercising using an MFA. The more benefits users obtain from the MFA, the greater their satisfaction [25]. Wang et al [9] also found that satisfaction was a partial mediator between CUI and PU. We found that the level of user satisfaction with an MFA was based not only on its meeting users' sports activity expectations but

XSL•FO
RenderX

also on the convenience and effectiveness of the features, the user interface, and the user experience that supported the user's sports activities:

> *I will continue to use Strava, because I am comfortable with Strava.* [Interviewee 14]
>
> *I will continue to use it, because in my opinion it is also effective and the appearance is not a hassle.* [Interviewee 10]
>
> *What makes me satisfied is the user interface, which is easy to use, and the user experience is simple.* [Interviewee 19]

Finally, the effect of satisfaction on CUI was confirmed in this study. The acceptance of H12 is in accordance with the studies by Bhattacherjee [38], Huang et al [15], Chiu et al [25], Wang et al [9], Cai et al [83], and Wu et al [66]. Wu et al [66] found that satisfied users are more likely to continue using an app because dissatisfied users can easily switch to other mHealth technologies. User satisfaction is an important determinant of the postadoption behavior of users of IT products and services [38]. This is supported by Chiu et al [25], who state that user satisfaction with the use of IT products and services is very important for fostering long-term use of IT. The main reason for user satisfaction with an MFA is that the features are complete and meet user needs, with the result that they come to depend on the MFA for their exercise routines:

> *Because I really like it and I have become very dependent on this application for sports. I don't want to exercise if there is no access to this application.* [Interviewee 9]
>
> *This application has fulfilled my daily needs.* [Interviewee 15]

Discussion

Principal Findings

The findings from this study extend previous research by examining the incentive system in MFAs [17,26] and the use of the ECT in the context of mHealth [25,33,38,62,66]. It also expands the understanding of the moderating effect of gender on incentive-based systems [68]. We found that MFAs and the incentives they offer have a strong influence on users' sports activity behaviors and on their intention to continue using the app. The results of this study indicate that the most influential feature of an MFA is the SM incentive feature. MFA users often do not feel like exercising or engaging in PA if the activity is not being tracked by their app. The SM feature was also found to have a greater impact on male users than on female users. This finding regarding gender differs from the results of a study by Yin et al [17], who stated that no gender trend was evident in the effectiveness of the SM feature. Furthermore, in contrast to the study by Yin et al [17], we found that SS had little effect on the PA of MFA users. The results of the qualitative interviews indicate that this is because the social circle of Indonesian MFA users is relatively small, and this small social circle affects the effectiveness of the SS feature.

MFA service providers should evaluate how different app features impact users of different genders to effectively motivate users to keep using their app in the long term. In addition, users feel more satisfied when their expectations regarding the use of an app are met. App developers can increase the PU of their MFA by using the users' social communities (eg, by creating social profile features, group exercises, sporting events organized by recognized organizations or communities, and personalized challenges or awards based on the user's sports activity history). App developers can improve the accuracy of the tracking feature, whether through a smartphone or a smartwatch, with the goal of providing users with more in-depth statistics and data. For user convenience, app providers should also develop tracking features that start automatically.

Limitations

The respondents to both the quantitative and qualitative studies were predominantly aged 17 to 25 years and female; thus, other moderating variables could be considered in a future study. The weak effect size for PU in Table 4 indicates that the differences or relationships between some variables were not significant. This suggests that there are other variables that might influence PU, which were not considered in this study. In future research, another variable that could be considered is PA. This could serve as a metric to determine whether using MFAs with specific incentives increases users' PA [17].

Conclusions

The results of the study show that SM, platform rewards, and EI can all influence the PU of MFAs. However, no relationship was found between SS and the PU of MFAs. Indonesians generally consider MFAs to be useful because these apps allow them to track their sports activities and also offer rewards and awards. The confirmation of a user's initial expectations also affects their perceptions of the usefulness of MFAs. PU and confirmation of expectations also affect user satisfaction with MFAs, which in turn influences the user's desire to continue using the MFA. In addition, gender was shown to influence user behavior when using MFAs. In future research, the scope of EI incentives could be expanded by considering financial reasons for exercising, other people's recommendations, and job demands, among other factors. We suggest considering tangible benefits as additional incentives to determine whether quantifiable benefits, such as assets or money, can increase a person's motivation to exercise.

Acknowledgments

This research was supported by a 2024 internal funding from the Faculty of Computer Science, Universitas Indonesia.

XSL•FO
RenderX

Data Availability

The data sets generated and analyzed during this study are not publicly available due to a lack of authorization to share these data.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Study questionnaire and interview questions.
[DOCX File , 21 KB - humanfactors_v11i1e50957_app1.docx]

References

1. Physical activity. World Health Organization. 2022. URL: https://www.who.int/news-room/fact-sheets/detail/physical-activity [accessed 2023-02-09]
2. The global status report on physical activity. World Health Organization. 2022. URL: https://tinyurl.com/mr3hnxnx [accessed 2023-02-09]
3. Global recommendations on physical activity for health. World Health Organization. URL: https://tinyurl.com/ymrcjnmk [accessed 2023-03-10]
4. Consumer sentiment in Indonesia during the coronavirus crisis. McKinsey & Company. 2022. URL: https://tinyurl.com/84bhd2ub [accessed 2023-02-09]
5. Fitness apps- Indonesia. Statista. 2023. URL: https://tinyurl.com/4ary4us8 [accessed 2023-03-09]
6. Edney SM, Olds TS, Ryan JC, Vandelanotte C, Plotnikoff RC, Curtis RG, et al. A social networking and gamified app to increase physical activity: cluster RCT. Am J Prev Med 2020 Feb;58(2):e51-e62 [FREE Full text] [doi: 10.1016/j.amepre.2019.09.009] [Medline: 31959326]
7. Mao X, Zhao X, Liu Y. mHealth app recommendation based on the prediction of suitable behavior change techniques. Decis Support Syst 2020 May;132:113248 [FREE Full text] [doi: 10.1016/j.dss.2020.113248]
8. Tu R, Hsieh P, Feng W. Walking for fun or for “likes”? The impacts of different gamification orientations of fitness apps on consumers’ physical activities. Sport Manag Rev 2021 Feb 03;22(5):682-693 [FREE Full text] [doi: 10.1016/j.smr.2018.10.005]
9. Wang Y, Wang Y, Greene B, Sun L. An analysis and evaluation of quality and behavioral change techniques among physical activity apps in China. Int J Med Inform 2020 Jan;133:104029 [FREE Full text] [doi: 10.1016/j.ijmedinf.2019.104029] [Medline: 31759245]
10. Baker RC, Kirschenbaum DS. Self-monitoring may be necessary for successful weight control. Behav Ther 1993;24(3):377-394 [FREE Full text] [doi: 10.1016/s0005-7894(05)80212-6]
11. Boutelle KN, Kirschenbaum DS. Further support for consistent self-monitoring as a vital component of successful weight control. Obes Res 1998 May;6(3):219-224 [FREE Full text] [doi: 10.1002/j.1550-8528.1998.tb00340.x] [Medline: 9618126]
12. Yan L, Tan Y. Good intentions, bad outcomes: the effects of mismatches between social support and health outcomes in an online weight loss community. Prod Oper Manag 2018 Jan 01;27(1):9-27 [FREE Full text] [doi: 10.1111/poms.12793]
13. Plangger K, Campbell C, Robson K, Montecchi M. Little rewards, big changes: using exercise analytics to motivate sustainable changes in physical activity. Inf Manag 2022 Jul;59(5):103216 [FREE Full text] [doi: 10.1016/j.im.2019.103216]
14. Uetake K, Yang N. Inspiration from the “biggest loser”: social interactions in a weight loss program. Mark Sci 2020 May;39(3):487-499 [FREE Full text] [doi: 10.1287/mksc.2018.1116]
15. Huang CK, Chen CD, Liu YT. To stay or not to stay? Discontinuance intention of gamification apps. Inform Technol Peopl 2019 Dec 02;32(6):1423-1445 [FREE Full text] [doi: 10.1108/itp-08-2017-0271]
16. Zhu Y, Wang R, Zeng R, Pu C. Does gender really matter? Exploring determinants behind consumers' intention to use contactless fitness services during the COVID-19 pandemic: a focus on health and fitness apps. Internet Res 2022 Mar 18;33(1):280-307 [FREE Full text] [doi: 10.1108/intr-07-2021-0454]
17. Yin Q, Li L, Yan Z, Guo C. Understanding the effects of self-peer-platform incentives on users' physical activity in mobile fitness apps: the role of gender. Inform Technol Peopl 2021 May 20;35(3):1054-1072 [FREE Full text] [doi: 10.1108/itp-10-2020-0705]
18. Miller AS, Cafazzo JA, Seto E. A game plan: gamification design principles in mHealth applications for chronic disease management. Health Informatics J 2016 Jun;22(2):184-193 [FREE Full text] [doi: 10.1177/1460458214537511] [Medline: 24986104]
19. de Zambotti M, Claudatos S, Inkelis S, Colrain IM, Baker FC. Evaluation of a consumer fitness-tracking device to assess sleep in adults. Chronobiol Int 2015;32(7):1024-1028 [FREE Full text] [doi: 10.3109/07420528.2015.1054395] [Medline: 26158542]

XSL•FO
RenderX

20. Sama PR, Eapen ZJ, Weinfurt KP, Shah BR, Schulman KA. An evaluation of mobile health application tools. JMIR Mhealth Uhealth 2014 May 01;2(2):e19 [FREE Full text] [doi: 10.2196/mhealth.3088] [Medline: 25099179]
21. Willms A, Rhodes RE, Liu S. Effects of mobile-based financial incentive interventions for adults at risk of developing hypertension: feasibility randomized controlled trial. JMIR Form Res 2023 Mar 24;7:e36562 [FREE Full text] [doi: 10.2196/36562] [Medline: 36961486]
22. Bo Y, Liu QB, Tong Y. The effects of adopting mobile health and fitness apps on hospital visits: quasi-experimental study. J Med Internet Res 2023 Jul 28;25:e45681 [FREE Full text] [doi: 10.2196/45681] [Medline: 37505809]
23. Yuan S, Ma W, Kanthawala S, Peng W. Keep using my health apps: discover users' perception of health and fitness apps with the UTAUT2 model. Telemed J E Health 2015 Sep;21(9):735-741 [FREE Full text] [doi: 10.1089/tmj.2014.0148] [Medline: 25919238]
24. Cho H, Chi C, Chiu W. Understanding sustained usage of health and fitness apps: incorporating the technology acceptance model with the investment model. Technol Soc 2020 Nov;63:101429 [FREE Full text] [doi: 10.1016/j.techsoc.2020.101429]
25. Chiu W, Cho H, Chi CG. Consumers' continuance intention to use fitness and health apps: an integration of the expectation–confirmation model and investment model. Inform Technol Peopl 2020 Jun 26;34(3):978-998 [FREE Full text] [doi: 10.1108/itp-09-2019-0463]
26. Huang G, Ren Y. Linking technological functions of fitness mobile apps with continuance usage among Chinese users: moderating role of exercise self-efficacy. Comput Human Behav 2020 Feb;103:151-160 [FREE Full text] [doi: 10.1016/j.chb.2019.09.013]
27. Teng X, Bao Z. Factors affecting users' stickiness of fitness apps: an empirical study based on the S-O-R perspective. Int J Sports Mark Spons 2022 Mar 30;23(4):823-840 [FREE Full text] [doi: 10.1108/ijsms-06-2021-0123]
28. Huang G. Does warm glow promote physical activity? Examining the relative effectiveness of self-benefiting versus other-benefiting incentives in motivating fitness app use by corporate sponsorship programs. Health Commun 2021 Nov;36(13):1719-1730 [FREE Full text] [doi: 10.1080/10410236.2020.1794548] [Medline: 32698621]
29. Pearson E, Prapavessis H, Higgins C, Petrella R, White L, Mitchell M. Adding team-based financial incentives to the Carrot Rewards physical activity app increases daily step count on a population scale: a 24-week matched case control study. Int J Behav Nutr Phys Act 2020 Nov 19;17(1):139-124 [FREE Full text] [doi: 10.1186/s12966-020-01043-1] [Medline: 33208166]
30. Lemola S, Gkiouleka A, Read B, Realo A, Walasek L, Tang NK, et al. Can a 'rewards-for-exercise app' increase physical activity, subjective well-being and sleep quality? An open-label single-arm trial among university staff with low to moderate physical activity levels. BMC Public Health 2021 Apr 23;21(1):782 [FREE Full text] [doi: 10.1186/s12889-021-10794-w] [Medline: 33892665]
31. Ryan RM, Deci EL. Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. Am Psychol 2000 Jan;55(1):68-78 [FREE Full text] [doi: 10.1037//0003-066x.55.1.68] [Medline: 11392867]
32. Teixeira PJ, Carraça EV, Markland D, Silva MN, Ryan RM. Exercise, physical activity, and self-determination theory: a systematic review. Int J Behav Nutr Phys Act 2012 Jun 22;9:78 [FREE Full text] [doi: 10.1186/1479-5868-9-78] [Medline: 22726453]
33. Li J, Liu X, Ma L, Zhang W. Users' intention to continue using social fitness-tracking apps: expectation confirmation theory and social comparison theory perspective. Inform Health Soc Care 2019 Sep;44(3):298-312 [FREE Full text] [doi: 10.1080/17538157.2018.1434179] [Medline: 29504821]
34. Bandura A. Social cognitive theory of self-regulation. Organ Behav Hum Decis Process 1991 Dec;50(2):248-287. [doi: 10.1016/0749-5978(91)90022-l]
35. Davis FD, Bagozzi RP, Warshaw PR. User acceptance of computer technology: a comparison of two theoretical models. Manag Sci 1989 Aug;35(8):982-1003 [FREE Full text] [doi: 10.1287/mnsc.35.8.982]
36. Barkley JE, Lepp A, Santo A, Glickman E, Dowdell B. The relationship between fitness app use and physical activity behavior is mediated by exercise identity. Comput Human Behav 2020 Jul;108:106313 [FREE Full text] [doi: 10.1016/j.chb.2020.106313]
37. Schoeppe S, Alley S, Van Lippevelde W, Bray NA, Williams SL, Duncan MJ, et al. Efficacy of interventions that use apps to improve diet, physical activity and sedentary behaviour: a systematic review. Int J Behav Nutr Phys Act 2016 Dec 07;13(1):127 [FREE Full text] [doi: 10.1186/s12966-016-0454-y] [Medline: 27927218]
38. Bhattacherjee A. Understanding information systems continuance: an expectation-confirmation model. MIS Q 2001 Sep;25(3):351 [FREE Full text] [doi: 10.2307/3250921]
39. Wang Q, Egelandsdal B, Amdam GV, Almli VL, Oostindjer M. Diet and physical activity apps: perceived effectiveness by app users. JMIR Mhealth Uhealth 2016 Apr 07;4(2):e33 [FREE Full text] [doi: 10.2196/mhealth.5114] [Medline: 27056639]
40. Chen Y, Pu P. Healthy together: exploring social incentives for mobile fitness applications. In: Proceedings of the 2nd International Symposium of Chinese CHI. 2014 Presented at: Chinese CHI '14; April 26-27, 2014; Toronto, ON p. 25-34 URL: https://doi.org/10.1145/2592235.2592240 [doi: 10.1145/2592235.2592240]
41. Bojd B, Song X, Tan Y, Yan X. Gamified challenges in online weight-loss communities. SSRN Journal. Preprint posted online April 25, 2018 2018 [FREE Full text] [doi: 10.2139/ssrn.3157331]

XSL•FO
RenderX

42. Payne HE, Lister C, West JH, Bernhardt JM. Behavioral functionality of mobile apps in health interventions: a systematic review of the literature. JMIR Mhealth Uhealth 2015 Feb 26;3(1):e20 [FREE Full text] [doi: 10.2196/mhealth.3335] [Medline: 25803705]
43. Bollók S, Takács J, Kalmár Z, Dobay B. External and internal sport motivations of young adults. Biomed Hum Kinet 2011;3:101-105 [FREE Full text] [doi: 10.2478/V10101-011-0022-5]
44. Koo S, Byon KK, Baker TA. Integrating event image, satisfaction, and behavioral intention: small-scale marathon event. Sport Mark Q 2014;23:59 [FREE Full text]
45. Low XT, Pyun DY. Consumers' perceived functions of and attitude toward corporate sponsors of small-scale amateur sporting events. Event Manag 2016 Jul 25;20(2):227-238 [FREE Full text] [doi: 10.3727/152599516x14643674421771]
46. Gabriel S, Gardner WL. Are there "his" and "hers" types of interdependence? The implications of gender differences in collective versus relational interdependence for affect, behavior, and cognition. J Pers Soc Psychol 1999 Sep;77(3):642-655 [FREE Full text] [doi: 10.1037//0022-3514.77.3.642] [Medline: 10510513]
47. Ryan RM, Williams GC, Heather H, Deci EL. Self-determination theory and physical activity: the dynamics of motivation in development and wellness. Hellenic J Psychol 2009;6(2):107-124 [FREE Full text]
48. Weisskopf WA. The duality of human existence: an essay on psychology and religion by David Bakan. Am J Sociol 1967;73(3):368-369 [FREE Full text] [doi: 10.1086/224499]
49. Spence JT, Helmreich RL. Masculinity and femininity: their psychological dimensions, correlates, and antecedents. Psychol Women Q 1978;4(1):144-146 [FREE Full text] [doi: 10.7560/764439]
50. Kimbrough AM, Guadagno RE, Muscanell NL, Dill J. Gender differences in mediated communication: women connect more than do men. Comput Human Behav 2013 May;29(3):896-900 [FREE Full text] [doi: 10.1016/j.chb.2012.12.005]
51. Sun Y, Lim KH, Jiang C, Peng JZ, Chen X. Do males and females think in the same way? An empirical investigation on the gender differences in web advertising evaluation. Comput Human Behav 2010 Nov;26(6):1614-1624 [FREE Full text] [doi: 10.1016/j.chb.2010.06.009]
52. Vilela AM, Nelson MR. Testing the selectivity hypothesis in cause-related marketing among generation Y: [when] does gender matter for short- and long-term persuasion? J Mark Comm 2013 Oct 11;22(1):18-35 [FREE Full text] [doi: 10.1080/13527266.2013.841272]
53. Sun H, Zhang P. The role of moderating factors in user technology acceptance. Int J Hum Comput Stud 2006 Feb;64(2):53-78 [FREE Full text] [doi: 10.1016/j.ijhcs.2005.04.013]
54. Leong LY, Ooi KB, Chong AY, Lin B. Modeling the stimulators of the behavioral intention to use mobile entertainment: does gender really matter? Comput Human Behav 2013 Sep;29(5):2109-2121 [FREE Full text] [doi: 10.1016/j.chb.2013.04.004]
55. Venkatesh V, Thong JY, Xu X. Consumer acceptance and use of information technology: extending the unified theory of acceptance and use of technology. MIS Q 2012;36(1):157-178. [doi: 10.2307/41410412]
56. Weman Josefsson K, Johnson U, Lindwall M. Short report: moderations in exercise motivation - gender and age moderates the relations of motivation quality and exercise behavior. Health Psychol Behav Med 2018 Apr 18;6(1):93-103 [FREE Full text] [doi: 10.1080/21642850.2018.1462706] [Medline: 34040823]
57. Albashrawi M, Motiwalla L. When IS success model meets UTAUT in a mobile banking context: a study of subjective and objective system usage. In: Proceedings of the 30th Annual Workshop of the Swedish Artificial Intelligence Society. 2017 Presented at: SAIS '17; May 15-16, 2017; Karlskrona, Sweden URL: https://aisel.aisnet.org/sais2017/1/
58. Susanto A, Chang Y, Ha Y. Determinants of continuance intention to use the smartphone banking services. Ind Manage Data Syst 2016;116:508-525 [FREE Full text] [doi: 10.1108/imds-05-2015-0195]
59. Oghuma AP, Libaque-Saenz CF, Wong SF, Chang Y. An expectation-confirmation model of continuance intention to use mobile instant messaging. Telemat Inform 2016 Feb;33(1):34-47 [FREE Full text] [doi: 10.1016/j.tele.2015.05.006]
60. Festinger L. A Theory of Cognitive Dissonance. Stanford, CA: Stanford University Press; 1957.
61. Bhattacherjee A. An empirical analysis of the antecedents of electronic commerce service continuance. Decis Support Syst 2001 Dec;32(2):201-214 [FREE Full text] [doi: 10.1016/s0167-9236(01)00111-7]
62. Hsu CL, Lin JC. What drives purchase intention for paid mobile apps? – An expectation confirmation model with perceived value. Electron Commer Res Appl 2015 Jan;14(1):46-57 [FREE Full text] [doi: 10.1016/j.elerap.2014.11.003]
63. Ambalov IA. A meta-analysis of IT continuance: an evaluation of the expectation-confirmation model. Telemat Inform 2018 Sep;35(6):1561-1571 [FREE Full text] [doi: 10.1016/j.tele.2018.03.016]
64. Dehghani M, Kim KJ, Dangelico RM. Will smartwatches last? factors contributing to intention to keep using smart wearable technology. Telemat Inform 2018 May;35(2):480-490 [FREE Full text] [doi: 10.1016/j.tele.2018.01.007]
65. Nascimento B, Oliveira T, Tam C. Wearable technology: what explains continuance intention in smartwatches? J Retail Consum Serv 2018 Jul;43:157-169 [FREE Full text] [doi: 10.1016/j.jretconser.2018.03.017]
66. Wu C, Zhou Y, Wang R, Huang S, Yuan Q. Understanding the mechanism between IT identity, IT mindfulness and mobile health technology continuance intention: an extended expectation confirmation model. Technol Forecast Soc Change 2022 Mar;176:121449 [FREE Full text] [doi: 10.1016/j.techfore.2021.121449]

XSL•FO
RenderX

67. Vinnikova A, Lu L, Wei J, Fang G, Yan J. The use of smartphone fitness applications: the role of self-efficacy and self-regulation. Int J Environ Res Public Health 2020 Oct 20;17(20):7639 [FREE Full text] [doi: 10.3390/ijerph17207639] [Medline: 33092090]
68. Sharma P, Sivakumaran B, Marshall R. Impulse buying and variety seeking: a trait-correlates perspective. J Bus Res 2010 Mar;63(3):276-283 [FREE Full text] [doi: 10.1016/j.jbusres.2009.03.013]
69. Sheikh Z, Yezheng L, Islam T, Hameed Z, Khan IU. Impact of social commerce constructs and social support on social commerce intentions. Inform Technol Peopl 2019 Feb 04;32(1):68-93 [FREE Full text] [doi: 10.1108/itp-04-2018-0195]
70. Ng JY, Ntoumanis N, Thøgersen-Ntoumani C, Deci EL, Ryan RM, Duda JL, et al. Self-determination theory applied to health contexts: a meta-analysis. Perspect Psychol Sci 2012 Jul;7(4):325-340 [FREE Full text] [doi: 10.1177/1745691612447309] [Medline: 26168470]
71. Hair Jr JF, Matthews LM, Matthews RL, Sarstedt M. PLS-SEM or CB-SEM: updated guidelines on which method to use. Int J Multivar Data Anal 2017;1(2):107 [FREE Full text] [doi: 10.1504/ijmda.2017.10008574]
72. Awang Z. SEM Made Simple: A Gentle Approach to Learning Structural Equation Modelling. Bangi, Malaysia: MPWS Rich Publication; 2015.
73. Stragier J, Vanden Abeele M, Mechant P, De Marez L. Understanding persistence in the use of online fitness communities: comparing novice and experienced users. Comput Human Behav 2016 Nov;64:34-42 [FREE Full text] [doi: 10.1016/j.chb.2016.06.013]
74. Sun M, Jiang LC. Linking social features of fitness apps with physical activity among Chinese users: evidence from self-reported and self-tracked behavioral data. Inf Process Manage 2022 Nov;59(6):103096 [FREE Full text] [doi: 10.1016/j.ipm.2022.103096]
75. Kim HM, Cho I, Kim M. Gamification aspects of fitness apps: implications of mHealth for physical activities. Int J Hum Comput Interact 2022 May 13;39(10):2076-2089 [FREE Full text] [doi: 10.1080/10447318.2022.2073322]
76. Goes PB, Guo C, Lin M. Do incentive hierarchies induce user effort? Evidence from an online knowledge exchange. Inf Syst Res 2016 Sep;27(3):497-516 [FREE Full text] [doi: 10.1287/isre.2016.0635]
77. Hamari J, Koivisto J. “Working out for likes”: an empirical study on social influence in exercise gamification. Comput Human Behav 2015 Sep;50:333-347 [FREE Full text] [doi: 10.1016/j.chb.2015.04.018]
78. Pappu R, Cornwell TB. Corporate sponsorship as an image platform: understanding the roles of relationship fit and sponsor–sponsee similarity. J Acad Mark Sci 2014 Feb 16;42(5):490-510 [FREE Full text] [doi: 10.1007/s11747-014-0373-x]
79. Yang J, Kanthawala S, Joo E, Kononova A. Can brand sponsorship increase download intention for mHealth apps? The role of issue relevance, brand involvement, and perceived app quality. J Promot Manag 2021 Dec 27;28(6):869-892 [FREE Full text] [doi: 10.1080/10496491.2021.2015513]
80. van Elburg FR, Klaver NS, Nieboer AP, Askari M. Gender differences regarding intention to use mHealth applications in the Dutch elderly population: a cross-sectional study. BMC Geriatr 2022 May 24;22(1):449 [FREE Full text] [doi: 10.1186/s12877-022-03130-3] [Medline: 35610577]
81. Forman EM, Manasse SM, Dallal DH, Crochiere RJ, Berry MP, Butryn ML, et al. Gender differences in the effect of gamification on weight loss during a daily, neurocognitive training program. Transl Behav Med 2021 Apr 26;11(4):1015-1022 [FREE Full text] [doi: 10.1093/tbm/ibaa091] [Medline: 32945860]
82. Brandts J, El Baroudi S, Huber SJ, Rott C. Gender differences in private and public goal setting. J Econ Behav Organ 2021 Dec;192:222-247 [FREE Full text] [doi: 10.1016/j.jebo.2021.09.012]
83. Cai J, Zhao Y, Sun J. Factors influencing fitness app users’ behavior in China. Int J Hum Comput Interact 2021 May 17;38(1):53-63 [FREE Full text] [doi: 10.1080/10447318.2021.1921483]

Abbreviations

CUI: continuance usage intention
ECT: expectation-confirmation theory
EI: external influence
MFA: mobile fitness app
mHealth: mobile health
PA: physical activity
PU: perceived usefulness
SDT: self-determination theory
SM: self-monitoring
SS: social support

XSL•FO
RenderX

Edited by K Adapa; submitted 18.07.23; peer-reviewed by M Mardiana, H Namba; comments to author 27.01.24; revised version received 05.02.24; accepted 29.04.24; published 05.06.24.

Please cite as:
Faizah A, Hardian AFA, Nandini RD, Handayani PW, Harahap NC
The Influence of Incentive-Based Mobile Fitness Apps on Users' Continuance Intention With Gender Moderation Effects: Quantitative and Qualitative Study
JMIR Hum Factors 2024;11:e50957
URL: https://humanfactors.jmir.org/2024/1/e50957
doi:10.2196/50957
PMID:38837199

©Aaya Faizah, Alifah Fatimah Azzahra Hardian, Rania Devina Nandini, Putu Wuri Handayani, Nabila Cyldea Harahap. Originally published in JMIR Human Factors (https://humanfactors.jmir.org), 05.06.2024. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on https://humanfactors.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Using a Smartwatch App to Understand Young Adult Substance Use: Mixed Methods Feasibility Study

Sahiti Kunchay[1], BTech; Ashley N Linden-Carmichael[2], PhD; Saeed Abdullah[1], PhD

[1]College of Information Sciences and Technology, The Pennsylvania State University, University Park, PA, United States

[2]Edna Bennett Pierce Prevention Research Center, The Pennsylvania State University, University Park, PA, United States

Corresponding Author:
Sahiti Kunchay, BTech
College of Information Sciences and Technology
The Pennsylvania State University
Westgate Building
University Park, PA, 16802
United States
Phone: 1 (814) 865 8947
Email: sahiti@psu.edu

Abstract

Background: Young adults in the United States exhibit some of the highest rates of substance use compared to other age groups. Heavy and frequent substance use can be associated with a host of acute and chronic health and mental health concerns. Recent advances in ubiquitous technologies have prompted interest and innovation in using technology-based data collection instruments to understand substance use and associated harms. Existing methods for collecting granular, real-world data primarily rely on the use of smartphones to study and understand substance use in young adults. Wearable devices, such as smartwatches, show significant potential as platforms for data collection in this domain but remain underused.

Objective: This study aims to describe the design and user evaluation of a smartwatch-based data collection app, which uses ecological momentary assessments to examine young adult substance use in daily life.

Methods: This study used a 2-phase iterative design and acceptability evaluation process with young adults (aged 18-25 y) reporting recent alcohol or cannabis use. In phase 1, participants (8/15, 53%) used the data collection app for 14 days on their Apple Watches to report their substance use patterns, social contexts of substance use, and psychosocial risk factors (eg, affect). After this 14-day deployment, the participants completed a user experience survey and a semistructured interview to record their perspectives and experiences of using the app. Formative feedback from this phase informed feature modification and refinement of the app. In phase 2, an additional cohort (7/15, 47%) used the modified app for 14 days and provided feedback through surveys and interviews conducted after the app use period.

Results: Analyses of overall app use patterns indicated high, consistent use of the app, with participants using the app for an average of 11.73 (SD 2.60) days out of 14 days of data collection. Participants reported 67 instances of substance use throughout the study, and our analysis indicates that participants were able to respond to ecological momentary assessment prompts in diverse temporal and situational contexts. Our findings from the user experience survey indicate that participants found the app usable and functional. Comparisons of app use metrics and user evaluation scores indicate that the iterative app design had a measurable and positive impact on users' experience. Qualitative data from the participant interviews highlighted the value of recording substance use patterns, low disruption to daily life, minimal overall burden, preference of platforms (smartphones vs smartwatches), and perspectives relating to privacy and app use in social contexts.

Conclusions: This study demonstrated the acceptability of using a smartwatch-based app to collect intensive, longitudinal substance use data among young adults. The findings document the utility of smartwatches as a novel platform to understand sensitive and often-stigmatized behaviors such as substance use with minimal burden.

(JMIR Hum Factors 2024;11:e50795) doi:10.2196/50795

KEYWORDS

smartwatches; substance use; ecological momentary assessment; mobile health; mHealth; human-centered design; feasibility studies; mobile phone

XSL•FO
RenderX

Introduction

Background

Young adults exhibit some of the highest rates of substance use across all age groups in the United States [1], including alcohol use (50.2% or 17.5 million people), cannabis use (25.9% or 9 million people), vaping nicotine (24% or 8.3 million people), and prescription psychotherapeutic drug use (7.3% or 2.5 million people). Substance use can be associated with significant long-term effects on individuals' health and well-being [2]. As such, there is an urgent need to understand, detect, and mitigate substance use among young adults.

There has been much prior work within the substance use domain in determining various psychological, social, and environmental factors that impact young adults' substance use behaviors. These studies [3-6] highlight the value of collecting mood, affect, situational, and social context data to assess how they affect substance use patterns in this population. In recent years, this domain has shifted from relying on cross-sectional surveys and retrospective data toward using ecological momentary assessments (EMAs) on a daily level to detect relevant within-person trends. With recent advances in ubiquitous technologies and the surge of interest in accessible and affordable health care, there has been an increasing focus on understanding substance use and associated consequences through technology-based solutions. Thus, in the aforementioned studies [3-6], smartphones have been the primary device of choice.

In addition to having adaptable interfaces that support EMAs, smartphones also have extensive sensors that show potential to unobtrusively detect substance use in young adult populations. Prior work in the ubiquitous-computing community has described apps that seek to collect and analyze data to predict drinking episodes. Several studies have investigated the efficacy of inferring alcohol use through a smartphone user's gait [7-9], as well as device use and movement features [10,11]. Smartphone sensors also exhibit potential in detecting cannabis use behaviors from users' gait using accelerometer and gyroscope data [12], as well as from a combination of time features and GPS, accelerometer, SMS text messaging, and smartphone logs [13].

Systems that are capable of capturing behaviors, experiences, and sensor data in real time provide researchers a deeper understanding of the various contexts in which young adults engage in substance use. Although smartphones have been successful in collecting such data and are thus widely used in substance use research, a recent review of EMA protocols determined that compliance for substance use–related EMAs deployed on participants' smartphones was lower than acceptable levels [14]. Hence, there is a need to explore novel interfaces and establish their utility in collecting granular substance use data with high compliance and low perceived burden.

Smartwatches offer a user experience that is distinct from that of a smartphone. The persistent, wearable nature of this device can enable users to observe cues (such as notifications, sounds, and vibrations) and perform quick interactions in diverse situations, such as when the smartphone is out of reach or an inconspicuous use of technology is required to minimize social disruption [15,16]. Moreover, smartwatches offer extensive health-sensing features that allow individuals to track and understand health behaviors. Thus, in recent years, there has been wide adoption of smartwatches: globally, approximately 202 million individuals own smartwatches [17], with 1 in 5 Americans using a smartwatch or fitness tracker [18]. This uptake of smartwatches by consumers has propelled researchers to investigate how smartwatches can be used as instruments of behavioral health studies. In fact, there have been several efforts to investigate whether illnesses and disorders could be recorded or managed through smartwatch-based tools such as those for managing attention-deficit/hyperactivity disorder [19] and posttraumatic stress disorder [20], aiding students with intellectual and developmental disabilities [21], assessing mobility among older adults [22], and managing chronic disorders [23]. In addition, there has also been interesting work in terms of detecting substance use behaviors, such as smoking, using these devices; for example, Skinner et al [24] used the accelerometer and gyroscope data in the Android Wear–based LG G Watch to detect signature hand movements of cigarette smoking.

Smartwatches and fitness trackers have met with resounding success in the health monitoring and self-management market [25,26]. Individuals use these devices to monitor and manage their fitness, sleep, mental health, and menstrual cycles through various apps. *Given their wide adoption for assessing health behaviors, especially by young adults [26], we argue that smartwatches may be well suited to understand substance use trends and patterns in this population*. In fact, Carreiro et al [27] highlight the significant potential yet underuse of wearables in combining detection and interventions for substance use. Importantly, the authors emphasized that wearable-smartphone combinations (such as smartwatches) are especially suitable for understanding and addressing substance use among adolescents and young adults. Recently, several studies pioneered the use of wearable sensors in understanding substance use and associated factors [28-30]. In these studies, participants noted several perceptions that suggested their preference for smartwatch-type interfaces over research-grade sensors for detecting and understanding substance use. Participants noted that these smartwatch interfaces were easy to integrate into their lives, offered various auxiliary features (such as screens, clock faces, and fitness-tracking capabilities), and drew minimal attention from strangers [28,29]. These aspects of smartwatches address many barriers that participants often face while using sensors and wearables in research studies. However, despite their potential and rapid uptake, these devices have rarely been used to assess substance use–related health behaviors in young adults.

Objectives

There is a critical need to better understand and assess substance use behaviors and trends in real-world settings, and this need has so far been addressed by using smartphones for collecting self-report and sensor data. However, the engagement and compliance rates of smartphone apps in this domain are less

XSL•FO
RenderX

than ideal, indicating a need to explore the suitability of other interfaces to collect such data. Therefore, this study aims to address this need by assessing the feasibility and acceptability of using smartwatches to collect EMA and sensor data to understand young adult substance use. Our use of smartwatches for this study is motivated by several reasons. First, in recent years, there has been wide adoption of commercially available smartwatches, specifically for health assessments and interventions. Second, smartwatches offer a novel user experience, built-in health sensor data capture, and popularity within young adult populations, thus offering the potential to collect richer, more granular data to understand young adults' substance use with minimal burden. Finally, existing research in the substance use domain suggests that smartwatches may be especially suitable for understanding young adults' substance use behaviors [27].

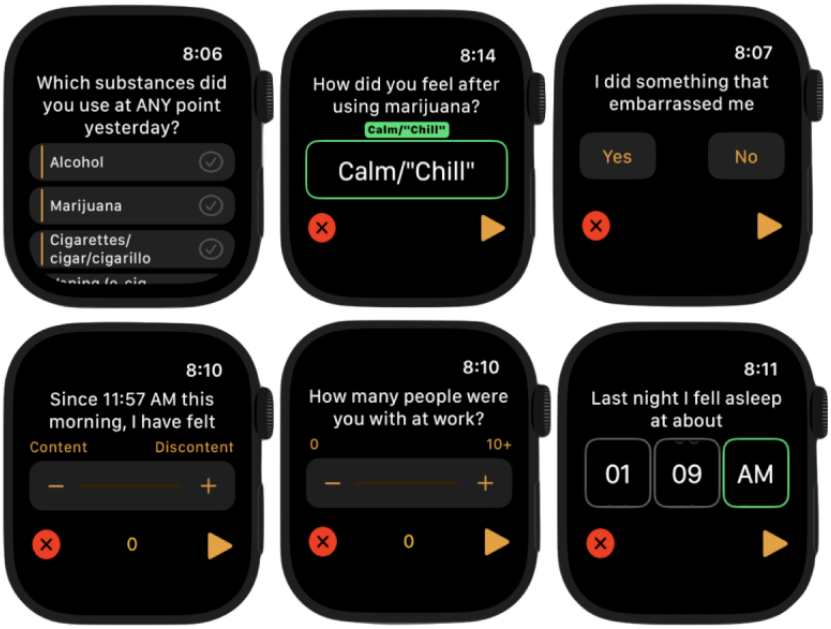
Methods

System Design and Development

Designing apps for smartwatches requires approaches and techniques that vary significantly from those required for typical smartphone app experiences. Smartwatch apps offer a seamless and intuitive experience when they are responsive; involve simple tasks; and make use of features that draw users to the device, such as haptic notifications, glanceable content, intuitive gestures, and a focused core functionality [31].

The primary requirement of the interface concerns ensuring that it provides an experience that results in highly granular and robust data collection, while keeping the perceived burden of interaction low. To address this challenge of high response rates and low study burden, we designed our questionnaire so that each question would take <5 seconds to answer. We expected that keeping the questions concise and interactions intuitive would help maintain low perceived burden and survey fatigue. Examples of these EMA components on the smartwatch (ie, an Apple Watch) are depicted in Figure 1. A companion app on the user's iPhone uploaded all user responses and sensor data to our database.

Figure 1. Screenshots of ecological momentary assessments on an Apple Watch interface depicting the variety of interface elements used to elicit responses from participants.



Iterative Application Design: Phase 1 and Phase 2

Our fundamental approach to app design and development was based on the principles of human-centered design, an approach that heavily incorporates users' experiences and perspectives throughout the design process. It is a nonlinear process that iterates continually between various stages of understanding users, defining the problem domain, generating ideas, prototyping or developing solutions, and testing. This process helps build mobile health (mHealth) systems that are usable, effective, and accessible [32,33].

The creation of the smartwatch app went through continual iterations of user evaluation and development to produce an experience that enables robust data collection while also ensuring that the app is easy to use, minimally invasive, and considerate of users' privacy and security concerns. Thus, we incorporated feedback from participants (8/15, 53%) in phase 1 of the feasibility study, so that participants (7/15, 47%) enrolled in the next phase were able to evaluate a refined app that provided a better user experience. During the initial rounds of testing and development, the app required confirmation from the database for each EMA question, causing a 1- to 2-second delay. This delay generated negative feedback from phase 1 participants. They reported that this delay between questions was frustrating and prompted them to assume that their responses were not recorded. To correct for this delay, we

XSL•FO
RenderX

eliminated the step of waiting for the database confirmation before moving to the next screen.

Participants and Procedures

Our methodology for the feasibility study was informed by two main objectives: (1) to fully capture participants' experience using the smartwatch app and their perspectives on its usability; and (2) to collect data that accurately reflect users' lived experiences with substance use, social contexts, affect, behaviors, and experiences. Hence, for this study, we used a mixed methods design with 3 key components: a 14-day in situ data collection period, where participants used the app to answer short EMAs regarding their behaviors and experiences; a poststudy survey that sought to quantitatively capture the usability of the app through various dimensions; and a semistructured interview that sought to capture more nuanced perspectives on participants' experiences with the app.

To be eligible, participants needed to be aged 18 to 25 years, report past-week alcohol or cannabis use, own and use both an iPhone (with iOS version 15 or newer) and an Apple Watch (with watchOS version 8 or newer) to deploy and use the smartwatch app, and be a current student at the local university.

Participants were recruited through convenience sampling, using study flyers posted on the university campus, social media posts, and the university's StudyFinder website. Potential participants were asked to email the study team if they were interested, after which they were sent a link to the screener survey as well as more details about the study. Informed consent to participate in the study was also obtained at this stage.

Eligible participants were immediately directed to a baseline survey in which they provided demographic information, typical substance use behavior, and technological use behaviors. The screener and baseline surveys were collected and managed using REDCap (Research Electronic Data Capture; Vanderbilt University) [34]. After completing the baseline survey, participants were scheduled for a web-based visit with the research staff who explained the research activities, guided them through app installation, and informed them about the compensation structure. Of the 25 eligible participants who completed the baseline survey, 15 (60%) scheduled and attended the web-based visit. After the completion of the 14-day data collection period, participants were requested via email to upload their HealthKit (Apple Inc) data, complete the usability survey, and schedule a second web-based visit for the semistructured interview.

Of the 15 participants who used the app, 12 (80%) completed all research activities. All study activities were conducted virtually between August 2021 and May 2022.

Measures

EMA Data

The types of data we collected from the user through the EMAs related to (1) mood and general affect [35]; (2) experiences of stress; (3) sleep duration; (4) types and amounts of substances used; (5) feelings of intoxication [36,37]; (6) substance use–related consequences [38]; and (7) social context, such as location and social environment. The questions for self-reports explored a wide range of constructs and were sourced from prior research and findings that established their validity and reliability [35-38]. All constructs used in this app were motivated by a wealth of research indicating various associations with substance use [5,39-45]. A full list of all aforementioned EMA items is included in Table S1 in Multimedia Appendix 1 [35-38]).

Participants were sent 5 survey prompts per day at 11 AM, 4 PM, 7 PM, 10 PM, and 1 AM, which were available only for specific time windows or *sessions* every day (11 AM-3 PM, 4 PM-6 PM, 7 PM-9 PM, 10 PM-midnight, and 1 AM-3 AM, respectively). A brief overview of the initial design and development of this app is provided in prior work [46]. For every item, participants had the option of skipping the question if they did not wish to respond.

In the 11 AM session, participants were asked about their experiences and behaviors that occurred at any time on the previous day, and these data were grouped as *prior day* data while analyzing responses. Participants were also asked (in all sessions) about their experiences and behaviors that occurred since their last response, and these data were categorized as *periodic* data during analysis (Table S1 in Multimedia Appendix 1).

Sensor Data

In addition to self-report questionnaires, we also collected sensor data: location (GPS), physical activity, and health data streams. The health data streams serve various purposes: physical exercise, exercise intensity, and the types of exercise are all factors that have significant benefits in reducing substance use, decreasing depression symptoms associated with substance use, and improving the abstinence rate among those using illicit substances [44,45]; sleep has a bidirectional relationship with substance use in young adults, with sleep patterns and duration being significant predictors of cigarette, alcohol, and cannabis use; and the type of substance use is a significant predictor of total sleep duration as well as sleep patterns (eg, weekend oversleep) [43]. Although the limited sample size in this study hinders us from assessing whether these data streams can be effectively leveraged to unobtrusively detect substance use behaviors, the feature is incorporated into the app to examine preliminary associations, as well as for use in future studies with an anticipated larger sample size.

User Experience Evaluation

For the usability survey, we used the System Usability Scale (SUS) [47] to assess the perceived usability of the Apple Watch app, and we used an adapted version of the Mobile Application Rating Scale: User Version (uMARS) [48] and various other items to assess the acceptability of the interface and the EMAs sourced from prior work [49,50]. Both the SUS and uMARS surveys have high reliability and validity and have been extensively used to evaluate digital systems and mHealth systems, respectively.

In the semistructured interview, we queried the participants on whether the app impacted their substance use or substance cravings; whether the app influenced their awareness of substance use patterns; whether they had any concerns about

XSL•FO
RenderX

using the app in various social contexts; and whether they had any privacy concerns regarding their substance use data, location data, or HealthKit data. All interviews were conducted via Zoom (Zoom Video Communications, Inc) and were recorded and transcribed using Zoom's live transcription service powered by Otter.ai. The interview script is provided in Textbox S1 in Multimedia Appendix 1.

Analysis

Only deidentified data were used during the analysis of app use data and interview data, blinding the authors to the identity of the participants while reviewing the results.

For our analysis, we focused on analyzing participants' EMA responses, app use patterns, and user perspectives to determine the feasibility and acceptability of the smartwatch app. To understand the effect of the iterative design improvements, we compared various measures between participants from both phases, treating them as separate groups during analysis. These findings are discussed in the *Results* section.

Ethical Considerations

All study activities and methods were approved by The Pennsylvania State University Institutional Review Board (17735) in the northeastern region of the United States in a state in which medical cannabis was legal, but recreational cannabis use was not legal at the time of data collection. A certificate of confidentiality was secured to protect participant responses concerning underage and illegal substance use behavior. All 15 participants provided informed consent before taking part in the study.

Participants were compensated for the study through Amazon gift cards and followed an established structure. Participants were compensated US $5 for completing the baseline survey, up to US $33 for the EMA data collection period, and US $10 for completing the user experience survey and semistructured interview. For the in-the-wild data collection period, participants were compensated US $2 per day if they completed both the 11 AM session and 1 other session during the day, but they were compensated only US $1 if they completed only the 11 AM session. Participants who did not complete the 11 AM session were not compensated for the day. If participants answered even 1 EMA during the 14-day period, they were compensated US $5.

Results

Quantity and Description of EMA Data Set

Participants ranged in age from 20 to 25 (mean 22.20, SD 1.86) years, and all were college students (undergraduate students: 10/15, 67% and graduate students: 5/15, 33%). Two-thirds (10/15, 67%) of the participants identified as female, while one-third (5/15, 33%) identified as male. Of the 15 participants, 5 (33%) identified as Asian, 1 (7%) as Black or African American, and 7 (47%) as White, while 1 (7%) participant preferred not to answer the question about race. Only 1 (7%) of the 15 participants identified as Hispanic or Latinx. Additional participant demographics are reported in Table S2 in Multimedia Appendix 1.

Overall, the 15 participants provided 4796 responses to EMA questions over 210 days. On average, the app collected 320 (SD 151; range 110-652) responses across all participants across all days of the study. Our data consisted of 45 prior-day (collected only at session 1) substance use reports, with a majority of reports mentioning alcohol use (alcohol: n=39, 87%; cannabis: n=12, 27%). We also collected 67 periodic substance use reports, which were reports collected in sessions 1, 2, 3, 4, or 5. Of these 67 periodic substance use reports, a majority included alcohol use, and a small portion included cannabis use, vape (e-cigarette or Juul e-cigarette) use, and cigarette use (reports of periodic alcohol use: n=49, 73%; reports of periodic cannabis use: n=13, 19%). Table 1 details all instances of substance use reported by the participants. Of the 15 participants, 3 (20%) did not report any substance use during the study.

XSL•FO
RenderX

Table 1. App use and substance use reports by participants.

Phase and participant	Total days participated (n=14), n (%)	Total sessions completed (n=70), n (%)	Days compliant (n=14), n (%)	Total EMAs[a] answered, n	Longest consecutive use of app (days; n=14), n (%)	Prior-day substance use reports (n=45), n (%)	Periodic substance use reports (n=67), n (%)
Phase 1							
P1	10 (71)	18 (26)	5 (36)	224	9 (64)	1 (2; alcohol)	1 (1; alcohol)
P2	8 (57)	27 (39)	7 (50)	271	8 (57)	2 (4; alcohol)	2 (3; alcohol)
P3	13 (93)	21 (30)	8 (57)	441	7 (50)	11 (24; alcohol: n=5, 45; cannabis: n=11, 100; vape: n=11, 100)	20 (30; alcohol: n=5, 25; cannabis: n=12, 60; vape: n=18, 90)
P4	13 (93)	25 (36)	6 (43)	259	10 (71)	1 (2; alcohol)	3 (4; alcohol)
P5	7 (50)	12 (17)	5 (36)	110	2 (14)	0 (0)	1 (1; other)
P6	14 (100)	55 (79)	14 (100)	559	14 (100)	1 (2; alcohol)	3 (4; alcohol)
P7	8 (57)	10 (14)	2 (14)	147	7 (50)	3 (7; alcohol)	4 (6; alcohol)
P8	14 (100)	36 (51)	12 (86)	424	14 (100)	4 (9; alcohol)	4 (6; alcohol)
Phase 2							
P9	14 (100)	36 (51)	10 (71)	371	14 (100)	0 (0)	0 (0)
P10	12 (86)	19 (27)	6 (43)	235	10 (71)	2 (4; alcohol: n=2, 100; cannabis: n=1, 50)	2 (3; alcohol: n=2, 100; cannabis: n=1, 50)
P11	14 (100)	28 (40)	13 (93)	380	14 (100)	9 (20; alcohol)	3 (4; alcohol: n=2, 67; cigarettes/cigar/cigarillo: n=1, 33)
P12	12 (86)	22 (31)	8 (57)	219	4 (29)	0 (0)	0 (0)
P13	14 (100)	26 (37)	10 (71)	289	14 (100)	0 (0)	0 (0)
P14	14 (100)	46 (66)	13 (93)	652	14 (100)	10 (22; alcohol)	22 (33; alcohol)
P15	9 (64)	21 (30)	4 (29)	215	9 (94)	1 (2; alcohol)	1 (1; alcohol)

[a]EMA: ecological momentary assessment.

When analyzing the intensity of substance use, we found that participants reported an average consumption of 3.44 (SD 3.09; min=1, max=≥10) alcoholic drinks, with more positive (mean 3.82, SD 2.05; range 0-6) than negative consequences related to alcohol use (mean 0.33, SD 0.66; range 0-3), while participants reporting prior-day cannabis use reported consuming an average of 8.92 (SD 2.23; min=3, max=≥10) hits, with an average of 2.667 (1.370; range 1-6) positive consequences and no negative consequences related to cannabis use.

Periodic substance use reports also included measures that asked participants to describe how they felt after consuming alcohol or cannabis. For alcohol use, the options provided were *buzzed*, *tipsy/happy*, *drunk*, and *wasted*. Most reports of alcohol use described participants feeling *buzzed* (12/28, 43%), followed by feeling *tipsy/happy* (9/28, 32%) and feeling *drunk* (7/28, 25%). For cannabis use, the options provided were *calm/chill*, *relaxed*, *high*, and *stoned*. Most reports of cannabis use described participants feeling *calm/chill* (3/7, 43%) or *high* (2/13, 29%). Cannabis use reports also included the manner in which the substance was consumed. A majority of responses reported cannabis use through pipes (7/13, 54%) or vapes (5/13, 38%).

Participants were also asked about various aspects of their health daily. In session 1, participants were asked about prior-day stress levels and sleep duration. In all sessions, participants were asked about their mood since the last response.

Of the 149 self-reports received for session 1, a total of 148 (99.3%) self-reports contained responses related to stress. In 96 (64.9%) of these 148 self-reports, participants reported that stressful events did not occur. When asked to rate their prior-day stress levels on a scale ranging from 1 to 100, on average, participants reported a stress level of 33.920 (SD 22.181; range 0-90). With respect to sleep, the app collected 147 self-reports, where participants were asked when they went to sleep the prior day and when they woke up on the current day. On average, participants reported 7.290 (SD 1.859; range 0-11.167) hours of sleep. Finally, participants were asked to report their mood through 8 bipolar items, which garnered 3167 self-reports.

Overall, the data collected through the app consisted of a broad range of substance use behaviors and experiences, as well as a variety of health behaviors. This suggests that *participants are willing and able to share substance use data through smartwatches*, along with a variety of measures that have

XSL•FO
RenderX

historically been associated with substance use in young adult populations.

App Use

We first examined how regularly participants used the app to answer EMAs. Of the 15 participants, 6 (40%) responded to at least 1 prompt on all 14 days of the study. Most participants (11/15, 73%) responded on ≥10 days. On average, participants provided data on 11.73 (SD 2.60) days out of the 14 days of the study. For participants completing all activities of the study, the average number of days participated was even higher: 12.24 (SD 2.14). Table 1 lists EMA completion details across each participant.

We had 403 sessions with ≥1 EMA response. On average, participants provided data for 26.80 (SD 12.15) sessions. We instructed participants to complete the first session every day along with at least 1 other session. Using these criteria, the overall compliance rate was 59% (8.2/14). On average, participants were compliant for 8.20 (SD 3.67) days out of the 14 days of the data collection period.

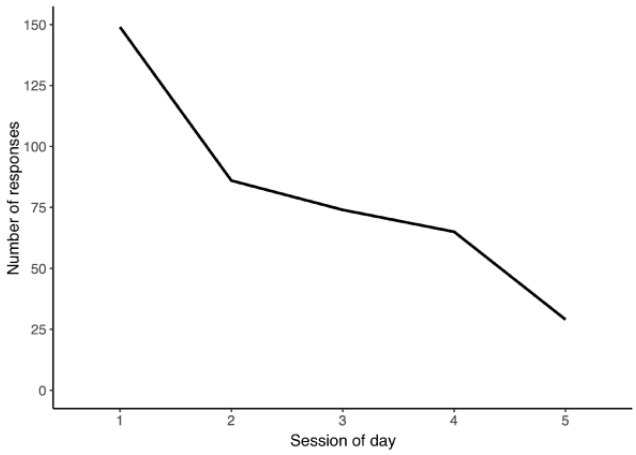
Finally, we also examined consecutive app use—the longest consecutive *streak* of days where participants used the app to provide responses. The longest streak was 14 days: 6 (40%) of the 15 participants used the app every day during the study. The average streak across all participants was 10.00 (SD 3.96) days, indicating sustained engagement with the app for a majority of the study duration.

Contextual Variations in App Use

In this part of our analysis, we wanted to determine whether there were certain times and contexts in which participants were less likely to respond to prompts than others. Toward this effort, we explored how app use patterns varied with time, substance use, and social environments.

In our data set, the response rate varied across sessions (Figure 2). Session 1 (11 AM-3 PM) had the most responses, and app use fell as the day progressed, with the lowest responses being collected during session 5 (1 AM-3 AM). To understand whether the session of day had a significant effect on whether the participant would respond, we used multilevel modeling (using the *lme4* package in R).

Figure 2. Variations in responses by session.



A null model allowed us to calculate the intraclass correlation coefficient (ICC) of whether a participant responded. The ICC was 0.135, which meant that only 13.5% of the variation in responding stemmed from between-person differences, which indicated that a large proportion of the variation arose due to within-person changes. Thus, a random intercept model was created by adding the session of day as a predictor. This model significantly explains more of the variance in participants' responses than the null model and hence is a better fit to the data (χ^2_4=181.6; *P*<.001). Using this model, we found that the session of day had a highly significant effect on whether the participant would respond. The odds and odds ratios calculated using this model indicated that the probability of a participant responding in session 1 was approximately 0.73. Compared to session 1, sessions 2, 3, 4, and 5 were respectively associated with a 76.47%, 82.29%, 85.88%, and 95.68% *decrease* in odds of a participant responding. In other words, *the probability of a participant answering in a particular session decreased significantly across the day*. Model details are described in Table S3 in Multimedia Appendix 1.

A random slope model did not significantly improve the fit of the model (χ^2_{14}=22.9; *P*=.06) and thus was not included for further analysis.

We used a similar method to understand whether the likelihood of reporting substance use varied across the day. The results of our null model calculated an ICC of 0.515, indicating that 51.5% of the variation in reporting substance use stemmed from between-person variances. The results from our model revealed that only session 5 had a significantly higher probability of participants reporting substance use compared to session 1 (estimate=1.70, SE 0.67; *P*=.01). The odds ratio for session 5 indicated that the odds of a participant reporting substance use in session 5 were approximately 5.49 times higher than the odds of a participant reporting substance use in session 1. This model proved to be a significantly better fit to the data than the null model (χ^2_4=11.3; *P*=.02), that is, participants were more likely to report substance use later in the day. Model details are described in Table S4 in Multimedia Appendix 1.

We also explored whether participants were able and likely to respond even when under the influence of substances.

XSL•FO
RenderX

Specifically, we analyzed how participants' responses differed after they reported substance use (compared with reports with no substance use). For this analysis, we used repeated measures correlations to determine within-individual association for paired or repeated measures data using the *rmcorr* package in R. We found no significant moment-level associations of substance use reports with responses in subsequent sessions, that is, whether a participant reported substance use in a specific session had no significant impact on their response to the first (r=0.02, 95% CI −0.08 to 0.13; P=.65), second (r=0.01, 95% CI −0.09 to 0.11; P=.83), third (r=0.03, 95% CI −0.07 to 0.13; P=.56), or fourth (r=−0.03, 95% CI −0.14 to 0.06; P=.46) session after the reported substance use. Random intercept multilevel models confirmed this result: reporting substance use in a specific session was not a significant predictor of whether a participant responded to the first, second, third, or fourth sessions after the session in question. Similarly, we saw no significant associations between social environments (people and places) and participants' likelihood of responding.

To summarize, our findings suggested that *participants were likely to respond to EMA prompts in a variety of social contexts and after consuming substances*. However, we found a time effect, where participants were more likely to respond to prompts earlier in the day.

Differences in Use Patterns Between Design Phases

To investigate whether the improvements made to the smartwatch app had any effect, we compared 5 metrics between phase1 and phase 2 participants with respect to the total number of EMAs answered, the total number of sessions completed, the total number of days participated, the total number of days compliant, and the longest consecutive use of the app. Before running the analysis, we used the Shapiro-Wilk test to check whether the metric values were distributed normally across the phases. The distributions of counts from participants in phase 2 for the total number of days participated (W=0.77; P=.03), the total number of EMAs answered (W=0.74; P=.02), and the longest consecutive use of the app (W=0.77; P=.02) were all significantly nonnormal. Thus, to test for differences between the phases for these 3 variables, we used a nonparametric test, the Wilcoxon rank sum test. For the remaining variables, we used the independent 2-tailed t test (the Welch 2-sample t test) to examine whether the differences were significant.

Our analysis showed that, on average, the total number of days participated among phase 1 participants (mean 10.88, SD 2.95) was lower than that among phase 2 participants (mean 12.71, SD 1.89); however, this difference was not significant (W=17; P=.21; r=−0.32). Similarly, the group means were higher for phase 2 participants compared to those for phase 1 participants in terms of the total number of EMAs answered (phase 1: mean 304.38, SD 155.84; phase 2: mean 337.29, SD 154.78; W=21; P=.86; r=−0.05), the total number of sessions completed (phase 1: mean 25.50, SD 14.55; phase 2: mean 28.29, SD 9.64; t_{12}=−0.44; P=.67), the total number of days compliant (phase 1: mean 7.38, SD 3.93; phase 2: mean 9.14, SD 3.39; t_{13}=−0.94; P=.37), and the longest consecutive use of the app (phase 1: mean 8.88, SD 3.94; phase 2: mean 11.29, SD 3.86; W=17; P=.21; r=−0.33), but none of the differences between the phases were significant.

Although we did not see statistically significant increases in app use metrics after improving the app, the systematically higher engagement in terms of days used, EMAs answered, days compliant, and longest consecutive use suggests that the changes were a step in the right direction.

User Evaluation

For our analysis of the user experience survey deployed after the participants finished their 14-day data collection period, we primarily focused on reporting various measures of usability, describing notable user perceptions, and comparing usability metrics between phase 1 and phase 2 participants to evaluate the effect of app improvements on overall user experience.

SUS Scores

Of the 15 participants who used the app, 12 (80%) completed all research activities. The average SUS score for all 12 participants was 63.54 (SD 18.78). As a comparison point, Bangor et al [51] found that the mean SUS score from 964 usability tests across various interface types was 70. However, a usability study of fitness trackers found that the average SUS score for an Apple Watch interface was 61.36 [52]. While slightly higher than average in terms of smartwatch interface, this score does provide the opportunity to understand pain points within the app. The mean score for each SUS measure is depicted in Table 2.

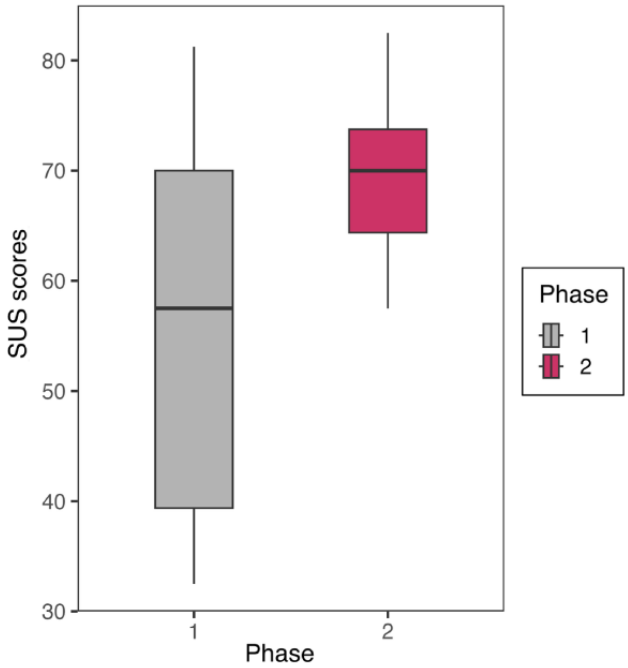
Table 2. Summary of itemized System Usability Scale (SUS) scores presented overall (combining the results of participants from both phases) and by study phase. Notably, phase 2 participants reported higher mean SUS scores than phase 1 participants, but this difference was not significant.

Measures	SUS items, mean (SD)										Total SUS score, mean (SD)
	1	2	3	4	5	6	7	8	9	10	
Overall	2.83 (1.11)	2.5 (1.17)	3 (1.28)	1.83 (0.94)	3.33 (1.30)	2.67 (1.44)	4.33 (0.49)	3 (1.35)	3.67 (1.07)	1.75 (0.96)	63.54 (18.78)
Phase 1	2.17 (1.17)	3 (1.26)	2.83 (0.98)	1.5 (0.84)	3.33 (1.63)	3.33 (1.63)	4.17 (0.40)	3.33 (1.63)	3.33 (1.21)	1.83 (1.16)	57.08 (22.77)
Phase 2	3.5 (0.55)	2 (0.89)	3.17 (1.60)	2.17 (0.98)	3.33 (1.03)	2 (0.89)	4.5 (0.55)	2.67 (1.03)	4 (0.89)	1.67 (0.82)	70 (12.55)

XSL•FO
RenderX

Before comparing overall mean SUS scores between the phases, we first used the Shapiro-Wilk test to check whether the SUS scores were distributed normally across the phases. The results of this test and an examination of skew and kurtosis values indicated that the SUS scores were distributed normally overall and by phase. Thus, to compare the means, we used the independent *t* test (the Welch 2-sample *t* test). *Although participants in phase 2 reported higher mean SUS scores than those in phase 1, this difference was not statistically significant* (phase 1: mean 57.08, SD 22.77; phase 2: mean 70.0, SD 12.55; t_8=−1.21; *P*=.26). A box plot depicting differences in the SUS scores between the phases is shown in Figure 3.

Figure 3. Box plot of System Usability Scale (SUS) scores by phase of study. Mean SUS scores were higher in phase 2, after we had made changes to the app to correct for delay issues. However, the difference in SUS scores between the phases was not statistically significant.



uMARS Scores

All uMARS items were rated on a scale ranging from 1 to 5. Both the average functionality and aesthetics scores for all 12 participants were 3.88 (SD 0.55) and 3.89 (SD 0.52) respectively, indicating that participants found both measures acceptable (mean ≥3).

In terms of overall mean scores, participants in phase 2 rated app functionality as *good* (mean ≥4), while those in phase 1 rated app functionality as *acceptable*. Participants from both phases rated aesthetics as *acceptable*. Mean functionality scores were higher in phase 2 (phase 1: mean functionality score=3.58, SD 0.61; phase 2: mean functionality score=4.17, SD 0.31; t_7=−2.11; *P*=.07), but mean aesthetics scores were slightly higher in phase 1 (phase 1: mean aesthetics score=3.94, SD 0.71; phase 2: mean aesthetics score=3.83, SD 0.28; t_7=−2.11; *P*=.07).

Notably, the scores for the performance domain of the functionality metric from phase 2 participants exceeded those from phase 1 participants (phase 1: mean performance score=2.33, SD 0.52; phase 2: mean performance score=4.17, SD 0.75; difference in scores between the 2 phases=1.84). Using the Wilcoxon rank sum test, *we found that phase 2 participants rated the app significantly higher on the performance scale than phase 1 participants* (W=1; *P*=.006; *r*=−0.79). There were no significant differences between the phases for the ease of use, navigation, and gestural design domains of uMARS functionality scores. Similarly, none of the aesthetics domains (layout, graphics, and visual appeal) had any significant differences in ratings between the phases.

EMA-Specific Participant Perceptions

Along with ratings of established usability scales, we also asked participants specific survey questions about EMA usability, touching upon the constructs of the ease of use, enjoyment, the speed of answering EMAs, EMA length, interruptibility, and notions of trust and privacy.

Most of the participants (8/12, 67%) either agreed or strongly agreed that the EMAs were easy to fill. However, there were discrepancies between the phases. Most participants in phase 2 (5/6, 83%) agreed or strongly agreed that the EMAs were easy to fill, while 3 (50%) of the 6 participants in phase 1 either disagreed or strongly disagreed with this statement. Similarly, all participants in phase 2 (6/6, 100%) agreed or strongly agreed that they were able to complete the EMAs quickly, while 4 (67%) of the 6 participants in phase 1 disagreed with this statement. Of the 12 participants, 8 (67%) did not think that the EMAs were too long, and among the 4 participants who did think so, a majority (n=3, 75%) were phase 1 participants.

To examine perceived burden and fatigue, participants were asked about the number of days after which they felt tired of answering the EMAs. On average, participants in phase 1 reported fatigue after 6.40 (SD 3.44; range 3-12) days, while participants in phase 2 reported fatigue after 9.17 (SD 3.13;

XSL•FO
RenderX

range 4-12) days. This difference among the phases was not statistically significant (t_8=−1.39; P=.20).

Overall, most of the participants felt that the app was acceptable and simple to use. A higher proportion of participants in phase 2 felt so compared to those from phase 1. None of the participant ratings of ease, speed, or fatigue were significantly different across the phases, but the higher ratings in phase 2 suggest that the app changes were a step in the right direction to improve app usability and address critical issues.

Interview Themes

Overview

In this subsection, we present our main findings from our analysis of the semistructured interviews that were conducted after the 14-day data collection period. For the analysis of the interview data, we used inductive thematic analysis to identify common themes using a qualitative interpretivist approach. The primary author conducted the initial analysis and then discussed the themes and codes with the other authors to ensure the validity of the primary findings and to reduce bias.

Overall, participants agreed that their experience using the smartwatch app to answer EMAs was easy, novel, and acceptable, but they also brought up certain key issues with app responsiveness and commented on the suitability of the smartwatch interface for this specific use case.

General App Perceptions

When asked about their overall experience using the app to answer EMAs, most of the participants shared that the app was generally easy to use. A participant recalled that using the app was quickly incorporated into their day, while noting that this was not disruptive to their routine:

> *I mean it kind of turned into, like, an everyday routine where, like I just expected it at certain times and I used to take time out and do it.* [P6]

This sentiment was echoed by another participant:

> *Since it is only like 3 to 5 minutes, I didn't think that's a very disruptive time point, like I could do it in between class or, if I was at dinner [or] lunch and I remembered, I'd typically do it then.* [P4]

Although not disruptive to their daily lives, this participant shared that using the app was different to how they normally used their smartwatch:

> *Disruptive? Not really. I don't normally look at my watch for more than a couple of seconds, so that was a little different, but overall it wasn't really that disruptive.* [P4]

Other participants also noted how completing the EMAs only took a few minutes (generally <2-3 min), unless lagging or responsiveness issues occurred.

Advantages and Challenges of Using Smartwatch Interfaces

Participants presented varied perspectives when it came to the elements of the smartwatch interface; for instance, some of the participants found advantages and preferred the fact that the smartwatch provided a small screen and a personal experience:

> *I think that's the one benefit that the watch did have, is that it's such a small screen that it's hard for anyone to, you know, look at what you're doing, on such a small screen. So the watch definitely had a benefit in kind of, like, protecting your privacy.* [P8]

By contrast, participants also noted that the small screen and the wearable experience presented a hindrance, with a participant sharing the challenges they faced in using the watch to answer EMAs:

> *Well, so for me, just having, just turning my arm and touching my watch is, I don't know if it's a range of motion thing, It's just not the most natural thing to me and so just having to be in this position, looking at my watch, touching stuff, I don't particularly like that.* [P3]

This participant indicated that a bigger screen would provide a more comfortable experience:

> *Just having a larger screen to be able to do everything on, I think it'd be a lot easier.* [P3]

Personal preferences factored greatly into how easy and intuitive participants found various aspects of the app experience. When asked about their perspectives on the various formats in which the questions were presented, such as sliders, checkboxes, and radio buttons, the responses were similarly varied. Some of the participants found all question formats easy to answer:

> *I think all of the formats were very straightforward and in terms of them, like, how they worked, I think they all worked just fine. There was no issue transitioning between the different formats.* [P8]

Others reported issues with the radio button and checkbox formats:

> *I think the multiple select got harder because just, like, being able to see all the options and then be able to click next on an Apple Watch screen [that] is kind of tiny, so in that sense, yes [was a difficult format to answer].* [P7]

Similarly, a participant faced challenges with the slider format:

> *Think the [slider] one, because I think I had to press, if I'm like, you know, perfectly energetic [on the MDMQ] then I had to go all the way plus plus plus plus plus, it was like, a lot of plusses. Other than that, the rest was great.* [P12]

App Responsiveness and Lagging Issues

Phase 1 participants frequently shared their experiences with recurring lag issues, noting that it lengthened app use time and caused disruptions and general frustration:

> *Overall, it was pretty easy and straightforward, but it did start to get frustrating switching between different survey prompts. It would get, like, frozen a lot. So I would click to go to the next prompt, I guess, and it would get frozen, so surveys that were supposed*

XSL•FO
RenderX

> *to take 2 minutes ended up taking upwards of 10 minutes because it would get frozen.* [P8]

> *A lot of the buttons weren't the most responsive, so you had to click them a couple of times before it would actually do anything. And sometimes I had to restart my watch because it just wouldn't have responded.* [P3]

As mentioned in the *Iterative Application Design: Phase 1 and Phase 2* subsubsection in the *Methods* section, we identified that this delay was caused by the data-uploading mechanism, which was corrected for phase 2. As a result, phase 2 participants did not report this frequent lag between questions in their interviews.

Comparisons to iPhone Platform

Several participants believed that having the option to answer the EMAs on both the iPhone and the Apple Watch would offer an easier and more seamless experience and provided the strengths of both devices to support this sentiment. A participant offered some context where such a system would prove useful for them:

> *So I know in the evening, sometimes, especially when I'm just, like, sitting on my couch, laying down, watching TV, I'll take my watch off to charge for the night, but I'll still have my phone with me. So, I'm not gonna get those alerts, if I'm not wearing my watch. So, it's nice to be able to switch, then, to the different interface on my phone, to use that.* [P14]

Another participant shared a similar perspective:

> *I guess, that [having the option to complete surveys on both devices] would be okay, because that way you can at least see the surveys, do on your watch, in case your phone is not in your hand, you still have the watch, you're wearing your watch, you have the option of both.* [P11]

By contrast, a participant shared a scenario where using a smartwatch would prove easier than using a smartphone:

> *Usually you have to open up the phone and then you have to take off your mask [to unlock it using facial recognition]. With the watch, you don't have to do anything, you just, you know, do with the 1 finger, which makes it a lot easier and better.* [P12]

Similarly, another participant noted as follows:

> *I check my watch more than I check my phone. I feel just time wise, and yeah, I feel like it'd be harder to use the phone. Like take my phone out and use it.* [P14]

However, most of the participants agreed that the larger screen size would provide a smoother experience while answering EMAs, with a participant sharing their perspective on how having a bigger screen would benefit their experience:

> *Just cause it's a little bigger, and you can just, like, do it on your phone while walking or something, and like on your watch you can't really do that.* [P10]

Self-Monitoring Substance Use

Several participants shared how using the app provided a valuable self-tracking experience that helped them think about their substance use patterns. Although this was not an intended use of the app, participants found a tangible benefit in keeping track of their substance use to answer the EMAs accurately:

> *It makes you cognizant of your usage, and it makes you cognizant, while looking at the questions, as to what, you know, could be impacted [by substance use].* [P11]

A participant shared how answering the EMAs helped them evaluate their substance use:

> *I think it just forced me to kind of analyze...like, I'd mainly only drink on the weekends, so it made me [think about] how I spend my weekends and how much I was using a substance in a specific time frame. So it made you kind of take a step back and analyze that, which is always, I think, shocking to people, how much or how little they may have been using a substance.* [P8]

Answering frequent EMAs about their substance use helped participants increase their awareness of their substance use patterns and behaviors.

A participant also shared an interesting perspective of how useful they found the self-monitoring aspect of using the app and how they experienced a lack of incentive to track their substance use after the 14-day data collection period ended:

> *I think, just being aware of, like, how many drinks I was consuming. Yeah, because if I don't have to track it, I don't remember how much I drink. So, because I was able to be like oh, like the next window is at 7 o'clock, like, my next notification at 7, like that. I've had to remember that I've had, you know, 2 drinks to put it in that notification.* [P14]

Other participants noted how they already mentally keep track of their substance use, but using the app made them reevaluate their use:

> *It definitely increased my awareness, but I felt like I already knew. If I had work or most of the school days, like, I won't be doing anything like that [substance use], but, more on the weekends. Like, oh, maybe I shouldn't do this tonight, or something like that.* [P7]

Participants used the EMAs to reflectively track their substance use. These interactions augmented their existing self-monitoring practices to periodically and contextually evaluate their substance use behaviors.

Use of the App in Social Settings

All participants reported that they used the app in public and social settings and were comfortable doing so; for instance, a participant shared how their friends felt when they saw the participant using the app:

> *Yeah, like I thought it was totally fine. All my friends knew I was taking, [and] like I didn't care that they*

XSL•FO
RenderX

> *knew. But when I was out at the bars, I was fine taking the surveys, and I don't know if other people knew that I was using my watch, or whatever. But all my friends knew, and they thought it was cool.* [P14]

Another participant also spoke about their use of the app in such settings:

> *Yeah, like, if I got the notification when I was at school, like in class or something, or like walking to class, I would take it then.* [P7]

This indicates that the app is able to effectively collect data in various social settings. This finding is especially meaningful, given the sensitive and often-stigmatized nature of the substance use data that the app collects. The convenience and comfort with which participants are able to share information indicates that using the smartwatch in this way is potentially unintrusive in various social contexts and environments.

Several participants offered insight into how they did not have concerns regarding privacy or security while interacting with the app and shared how the smartwatch platform helped in this aspect:

> *No [I did not feel uncomfortable using the app in public or social settings]. I mean, the watch screen is so small, I don't even think anyone realized what I was doing, that I'm on it.* [P7]

The small screen of the watch ensured that the participants' activities while using the app remained private from their peers and other people in their vicinity and thus helped their perception of the security of their data.

Discussion

App Feasibility and Acceptability

Overall, the app collected 4796 responses to EMA questions from 15 participants over the course of a 2-week-long study. Participants demonstrated high and consistent use of the app, responding on an average of 11.73 (SD 2.60) days and consistently using the app for an average of 10 (SD 3.96) days. Our analysis of app use patterns indicates that participants respond in a variety of contexts: after they consume substances and among different social contexts. The interview data supported these findings: participants were able to quickly incorporate using the app into their daily life and easily provide substance use data, and they were comfortable using the app in diverse social settings. Together, these findings demonstrate that it is indeed feasible to use a smartwatch app to collect substance use data.

With respect to app use, the decrease in participants' responses across the day was an interesting finding. We speculate that the higher response rate in session 1 might be due to the longer availability compared to other sessions (4 h vs 2 h). Participants were also specifically asked to complete session 1 each day and were compensated accordingly. However, our findings also indicate that participants were more likely to report substance use at night, in session 5 (1 AM-3 AM), than in session 1 (11 AM-3 PM), which coincides with substance use patterns among young adults. Together, these results suggest that there are certain time periods that may be better suited to obtaining specific insights into substance use behaviors. Morning and noon may be suitable periods to understand prior-day substance use behaviors, mood, and experiences, while late night might be better suited to understand evening drinking behaviors. As such, there is an opportunity to develop better informed and less burdensome methods for collecting substance use data. Future work should try to replicate our findings regarding the temporal variation of EMA completion rates for substance use.

In terms of user evaluations, the average SUS score for the 12 participants who completed the survey was 63.54 (SD 18.78). Participants in phase 2 reported higher mean SUS scores than those in phase 1 (phase 1: mean 57.08, phase 2: mean 70.00). For context, an SUS score of 70.00 is considered average and acceptable, but it is to be noted that this subjective qualification of SUS scores does not consider smartwatch interfaces. If we factor the interface into our assessment of participants' SUS scores, we can estimate that overall and in phase 2, participants rated the app above average in usability. Furthermore, in terms of mean uMARS scores, participants from phase 2 rated app functionality as good (mean ≥4), while those from phase 1 rated app functionality as acceptable. Although not significant, these findings suggest that the performance improvements to the app had a large and measurable impact on participants' perceptions of usability. Indeed, the improvement also had an impact on app use: on average, the total number of days participated, the total number of EMAs answered, the total number of sessions completed, the total number of days compliant, and the longest consecutive use of the app were all higher among phase 2 participants than among phase 1 participants.

These findings not only establish the user acceptability of smartwatches to collect substance use data but also indicate that app performance, specifically responsiveness to user inputs, is critical for user acceptance. Given the limited computational capability of smartwatches, it is particularly important to aim for responsive design by default. Modifications to improve app responsivity resulted in better perceived usability and user satisfaction, along with systematically higher user evaluation scores and app use metrics. Thus, supporting quick, responsive interactions is a critical consideration when designing EMAs for smartwatches. Researchers and practitioners interested in using these devices as platforms for intensive data collection must focus on efficient, quick, and simple interactions to ensure sustained use as well as acceptable compliance and response rates.

Smartwatches and Substance Use

Our data consisted of 45 prior-day and 67 periodic substance use reports which contain alcohol, cannabis, cigarette/cigar/cigarillo, and e-cigarette/vape use data. Participants were able to share data on a range of variables associated with substance use through the smartwatch.

Furthermore, our interview data highlighted a key benefit that participants found through regularly using the app: tracking and reflecting on their substance use. Using the app to provide substance use data encouraged participants to contemplate on their substance use by requiring them to recollect aspects of their use (when, how much, with whom, etc). Even without a

XSL•FO
RenderX

feature that displays the patterns of use, participants noted how the task of recollecting and entering substance use data helped to make them more aware of patterns within their substance use as well as cognizant of the contexts in which they consume various substances. While the benefits of self-monitoring substance use are not limited to the smartwatch interface, it is promising that a smartwatch app is able to successfully promote such experiences.

An aspect of the smartwatch interface that might have helped participants share substance use data confidently is the privacy that it affords through a smaller, more discreet screen. Participants reported how they felt comfortable using the app in social and public settings, saying that the small screen ensured that others in their vicinity would not be able to discern what the participants were doing on their smartwatch. Nevertheless, some participants thought that a larger screen, such as a smartphone screen, might be useful in certain contexts. Participants also noted that some question formats, such as those that require scrolling, are harder to complete on a small screen. Importantly, participants preferred having the option to complete a survey on a smartphone or a smartwatch, depending on what is most convenient at a given time and place. Future studies using smartwatches for health assessments and interventions should ensure that the proposed systems can work comfortably across diverse contexts. One way to accomplish this is by supporting interchangeable use of the app on different devices: smartphones *as well as* smartwatches. Users can then choose which device is most appropriate for their current activity and social environment and use the app correspondingly.

On the whole, our analysis of app use, surveys, and interview data indicate the feasibility and acceptability of using smartwatches in this domain, demonstrating that users are able and willing to use a smartwatch to share substance use data. Participants shared data on a range of substances, experiences, and behaviors and identified aspects of the smartwatch interface that enabled them to do so comfortably. Participants found that comprehensively self-monitoring their substance use through the app was a useful and important feature. Our findings also provide insight into which aspects of the smartwatch interface elicit responses as well as those that do not: while the small screen affords users privacy while relaying sensitive information such as substance use data, it can provide challenges for certain EMA formats and in certain contexts.

Limitations

This study has a number of limitations that are important to discuss, given their potential impact on our findings. First, the study had a small sample size, which we considered to be acceptable, given that the goal of the study was to establish the feasibility and acceptability of a smartwatch-based app for collecting substance use data. However, we acknowledge that the reported findings may not be generalizable to the larger population of young adults who consume substances. Reproducing this study with a larger, more diverse sample can offer a wider perspective on the use of smartwatch-based apps to collect longitudinal, intensive data in this domain. The findings concerning significance should be interpreted with caution, given our small sample size and unstable estimates. Furthermore, participants were already smartwatch owners, which might have had an impact on the perceptions of usability. Thus, understanding the perspectives of novice smartwatch users using the app can help us investigate the effect of novelty on user experience. Finally, most participants in our sample were not binge drinkers or did not exhibit high-intensity substance use. The user experience of the app might differ with participants and circumstances that arise from heavy or hazardous substance use behaviors that are not adequately represented in our sample. Future work focusing on users who exhibit such patterns of substance use can help build more robust systems that cater to a wider range of people who use substances.

Next, we detail limitations associated with our app. Developing data collection apps on the Apple system has the constraint of being platform dependent (limiting the devices on which the apps can be deployed); however, developing for a single ecosystem was the first step in testing the general feasibility of a smartwatch-based data collection app. Implementing the data collection app on multiple ecosystems and running studies with various devices and apps was outside the scope of this study. However, our design and development process focused considerably on creating a reproducible and well-documented codebase so that cross-platform or platform-agnostic implementation can be achieved at a later stage.

Conclusions and Future Work

In recent years, there has been wide adoption of smartwatches for health assessments and interventions. This paper focuses on ascertaining the feasibility and acceptability of using a smartwatch app to collect substance use data from young adults. Our data indicate that it is feasible and acceptable to use smartwatches to collect data about sensitive and stigmatized behaviors, including substance use. On the basis of these findings, we also discuss considerations for future smartwatch apps for health and well-being data collection. These findings have important implications for researchers aiming to leverage smartwatches as an mHealth platform for effective assessments and interventions. In the future, we plan to conduct a larger study, with a randomized between-participants experiment design, to compare app use and user perceptions between smartphones and smartwatches. This future study will help us understand which device results in better compliance, better engagement, and lower perceived burden within the context of substance use data collection. We also intend to use the health sensor data from this larger study to explore whether they can be used to unobtrusively detect substance use or associated behaviors. Finally, we aim to incorporate analyses such as the impact of battery life on app use to gain a nuanced understanding of how smartwatch capabilities impact user experience.

XSL•FO
RenderX

Acknowledgments

The research reported in this paper was supported by the National Institute on Alcohol Abuse and Alcoholism of the National Institutes of Health (K01 AA026854) and a seed grant from the College of Information Sciences and Technology at The Pennsylvania State University.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Study materials, participant demographics, and multilevel model details.
[DOCX File , 27 KB - humanfactors_v11i1e50795_app1.docx]

References

1. Substance Abuse and Mental Health Services Administration. Key substance use and mental health indicators in the United States: results from the 2022 national survey on drug use and health. Center for Behavioral Health Statistics and Quality, Substance Abuse and Mental Health Services Administration. 2023. URL: https://www.samhsa.gov/data/report/2022-nsduh-annual-national-report [accessed 2024-04-09]
2. Substance misuse prevention for young adults. Substance Abuse and Mental Health Services Administration. 2019. URL: https://store.samhsa.gov/product/substance-misuse-prevention-young-adults/pep19-pl-guide-1 [accessed 2023-05-27]
3. Braitman AL, Linden-Carmichael AN, Henson JM. Protective behavioral strategies as a context-specific mediator: a multilevel examination of within- and between-person associations of daily drinking. Exp Clin Psychopharmacol 2017 Jun;25(3):141-155 [FREE Full text] [doi: 10.1037/pha0000111] [Medline: 28240925]
4. Linden-Carmichael AN, Allen HK, Lanza ST. The socio-environmental context of simultaneous alcohol and marijuana use among young adults: examining day-level associations. Drug Alcohol Rev 2021 May 13;40(4):647-657 [FREE Full text] [doi: 10.1111/dar.13213] [Medline: 33188538]
5. Lipperman-Kreda S, Paschall MJ, Robert FS, Morrison CN. Places and social contexts associated with simultaneous use of alcohol, tobacco and marijuana among young adults. Drug Alcohol Rev 2018 Feb 19;37(2):188-195 [FREE Full text] [doi: 10.1111/dar.12537] [Medline: 28422352]
6. Terry-McElrath YM, O'Malley PM, Johnston LD. Alcohol and marijuana use patterns associated with unsafe driving among U.S. high school seniors: high use frequency, concurrent use, and simultaneous use. J Stud Alcohol Drugs 2014 May;75(3):378-389 [FREE Full text] [doi: 10.15288/jsad.2014.75.378] [Medline: 24766749]
7. Arnold Z, Larose D, Agu E. Smartphone inference of alcohol consumption levels from gait. In: Proceedings of the International Conference on Healthcare Informatics. 2015 Presented at: ICHI 2015; October 21-23, 2015; Dallas, TX. [doi: 10.1109/ichi.2015.59]
8. Kao HL, Ho BJ, Lin AC, Chu HH. Phone-based gait analysis to detect alcohol usage. In: Proceedings of the 2012 ACM Conference on Ubiquitous Computing. 2012 Presented at: Ubicomp '12; September 5-8, 2012; Pittsburgh, PA. [doi: 10.1145/2370216.2370354]
9. Suffoletto B, Gharani P, Chung T, Karimi H. Using phone sensors and an artificial neural network to detect gait changes during drinking episodes in the natural environment. Gait Posture 2018 Feb;60:116-121 [FREE Full text] [doi: 10.1016/j.gaitpost.2017.11.019] [Medline: 29179052]
10. Bae S, Ferreira D, Suffoletto B, Puyana JC, Kurtz R, Chung T, et al. Detecting drinking episodes in young adults using smartphone-based sensors. Proc ACM Interact Mob Wearable Ubiquitous Technol 2017 Jun 30;1(2):1-36 [FREE Full text] [doi: 10.1145/3090051] [Medline: 35146236]
11. Bae S, Chung T, Ferreira D, Dey AK, Suffoletto B. Mobile phone sensors and supervised machine learning to identify alcohol use events in young adults: implications for just-in-time adaptive interventions. Addict Behav 2018 Aug;83:42-47 [FREE Full text] [doi: 10.1016/j.addbeh.2017.11.039] [Medline: 29217132]
12. Li R, Agu E, Balakrishnan G, Herman D, Abrantes A, Stein M, et al. WeedGait: unobtrusive smartphone sensing of marijuana-induced gait impairment by fusing gait cycle segmentation and neural networks. In: Proceedings of the IEEE Healthcare Innovations and Point of Care Technologies. 2019 Presented at: HI-POCT 2019; November 20-22, 2019; Bethesda, MD. [doi: 10.1109/hi-poct45284.2019.8962787]
13. Bae SW, Chung T, Islam R, Suffoletto B, Du J, Jang S, et al. Mobile phone sensor-based detection of subjective cannabis intoxication in young adults: a feasibility study in real-world settings. Drug Alcohol Depend 2021 Nov 01;228:108972 [FREE Full text] [doi: 10.1016/j.drugalcdep.2021.108972] [Medline: 34530315]
14. Jones A, Remmerswaal D, Verveer I, Robinson E, Franken IH, Wen CK, et al. Compliance with ecological momentary assessment protocols in substance users: a meta-analysis. Addiction 2019 Apr;114(4):609-619 [FREE Full text] [doi: 10.1111/add.14503] [Medline: 30461120]

XSL•FO
RenderX

15. Cecchinato ME, Cox AL, Bird J. Smartwatches: the good, the bad and the ugly? In: Proceedings of the 33rd Annual ACM Conference Extended Abstracts on Human Factors in Computing Systems. 2015 Presented at: CHI EA '15; April 18-23, 2015; Seoul, Republic of Korea. [doi: 10.1145/2702613.2732837]
16. Visuri A, Sarsenbayeva Z, van Berkel N, Goncalves J, Rawassizadeh R, Kostakos V, et al. Quantifying sources and types of smartwatch usage sessions. In: Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems. 2017 Presented at: CHI '17; May 6-11, 2017; Denver, CO. [doi: 10.1145/3025453.3025817]
17. Number of users of smartwatches worldwide from 2019 to 2028. Statista. 2023 Nov. URL: https://www.statista.com/forecasts/1314339/worldwideusers-of-smartwatches [accessed 2024-04-09]
18. Vogels EA. About one-in-five Americans use a smart watch or fitness tracker. Pew Research Center. 2020 Jan 09. URL: https://www.pewresearch.org/short-reads/2020/01/09/about-one-in-five-americans-use-a-smart-watch-or-fitness-tracker/ [accessed 2023-05-27]
19. Dibia V. FOQUS: a smartwatch application for individuals with ADHD and mental health challenges. In: Proceedings of the 18th International ACM SIGACCESS Conference on Computers and Accessibility. 2016 Presented at: ASSETS '16; October 23-26, 2016; Reno, NV. [doi: 10.1145/2982142.2982207]
20. Larsen JE, Eskelund K, Christiansen TB. Active self-tracking of subjective experience with a one-button wearable: a case study in military PTSD. arXiv. Preprint posted online on March 9, 2017 [FREE Full text]
21. Zheng H, Motti VG. WeLi: a smartwatch application to assist students with intellectual and developmental disabilities. In: Proceedings of the 19th International ACM SIGACCESS Conference on Computers and Accessibility. 2017 Presented at: ASSETS '17; October 20-November 1, 2017; Baltimore, MD. [doi: 10.1145/3132525.3134770]
22. Rouzaud Laborde C, Cenko E, Mardini MT, Nerella S, Kheirkhahan M, Ranka S, et al. Satisfaction, usability, and compliance with the use of smartwatches for ecological momentary assessment of knee osteoarthritis symptoms in older adults: usability study. JMIR Aging 2021 Jul 14;4(3):e24553 [FREE Full text] [doi: 10.2196/24553] [Medline: 34259638]
23. Schickler M, Pryss R, Reichert M, Heinzelmann M, Schobel J, Langguth B, et al. Using wearables in the context of chronic disorders: results of a pre-study. In: Proceedings of the IEEE 29th International Symposium on Computer-Based Medical Systems. 2016 Presented at: CBMS 2016; June 20-24, 2016; Belfast and Dublin, Ireland. [doi: 10.1109/cbms.2016.40]
24. Skinner AL, Stone CJ, Doughty H, Munafò MR. StopWatch: the preliminary evaluation of a smartwatch-based system for passive detection of cigarette smoking. Nicotine Tob Res 2019 Jan 04;21(2):257-261 [FREE Full text] [doi: 10.1093/ntr/nty008] [Medline: 29373720]
25. Dhingra LS, Aminorroaya A, Oikonomou EK, Nargesi AA, Wilson FP, Krumholz HM, et al. Use of wearable devices in individuals with or at risk for cardiovascular disease in the US, 2019 to 2020. JAMA Netw Open 2023 Jun 01;6(6):e2316634 [FREE Full text] [doi: 10.1001/jamanetworkopen.2023.16634] [Medline: 37285157]
26. DeMarco J. Nearly 70% of Americans would wear a fitness tracker/smartwatch for discounted health insurance. Value Penguin. 2022 Apr 18. URL: https://www.valuepenguin.com/fitness-tracker-smartwatch-health-survey [accessed 2023-05-27]
27. Carreiro S, Chai PR, Carey J, Lai J, Smelson D, Boyer EW. mHealth for the detection and intervention in adolescent and young adult substance use disorder. Curr Addict Rep 2018 Jun;5(2):110-119 [FREE Full text] [doi: 10.1007/s40429-018-0192-0] [Medline: 30148037]
28. Chapman BP, Lucey E, Boyer EW, Babu KM, Smelson D, Carreiro S. Perceptions on wearable sensor-based interventions for monitoring of opioid therapy: a qualitative study. Front Digit Health 2022 Oct 21;4:969642 [FREE Full text] [doi: 10.3389/fdgth.2022.969642] [Medline: 36339518]
29. Carreiro S, Chintha KK, Shrestha S, Chapman B, Smelson D, Indic P. Wearable sensor-based detection of stress and craving in patients during treatment for substance use disorder: a mixed methods pilot study. Drug Alcohol Depend 2020 Apr 01;209:107929 [FREE Full text] [doi: 10.1016/j.drugalcdep.2020.107929] [Medline: 32193048]
30. Holtyn AF, Bosworth E, Marsch LA, McLeman B, Meier A, Saunders EC, et al. Towards detecting cocaine use using smartwatches in the NIDA clinical trials network: design, rationale, and methodology. Contemp Clin Trials Commun 2019 Jun 04;15:100392 [FREE Full text] [doi: 10.1016/j.conctc.2019.100392] [Medline: 31245651]
31. Designing for watchOS. Apple Inc. URL: https://developer.apple.com/design/human-interface-guidelines/designing-for-watchos [accessed 2023-05-29]
32. McCurdie T, Taneva S, Casselman M, Yeung M, McDaniel C, Ho W, et al. mHealth consumer apps: the case for user-centered design. Biomed Instrum Technol 2012;Suppl:49-56. [doi: 10.2345/0899-8205-46.s2.49] [Medline: 23039777]
33. Monitoring and evaluating digital health interventions: a practical guide to conducting research and assessment. World Health Organization. 2016 Dec 12. URL: https://www.who.int/publications/i/item/9789241511766 [accessed 2024-06-10]
34. Harris PA, Taylor R, Thielke R, Payne J, Gonzalez N, Conde JG. Research electronic data capture (REDCap)--a metadata-driven methodology and workflow process for providing translational research informatics support. J Biomed Inform 2009 Apr;42(2):377-381 [FREE Full text] [doi: 10.1016/j.jbi.2008.08.010] [Medline: 18929686]
35. Wilhelm P, Schoebi D. Assessing mood in daily life. Eur J Psychol Assessment 2007 Nov 05;23(4). [doi: 10.1027/1015-5759.23.4.258]
36. Cloutier RM, Calhoun BH, Lanza ST, Linden-Carmichael AN. Assessing subjective cannabis effects in daily life with contemporary young adult language. Drug Alcohol Depend 2022 Jan 01;230:109205 [FREE Full text] [doi: 10.1016/j.drugalcdep.2021.109205] [Medline: 34890928]

XSL•FO
RenderX

37. Linden-Carmichael AN, Calhoun BH. Measuring subjective alcohol effects in daily life using contemporary young adult language. Exp Clin Psychopharmacol 2022 Apr;30(2):151-158 [FREE Full text] [doi: 10.1037/pha0000447] [Medline: 33591776]
38. Lee CM, Cronce JM, Baldwin SA, Fairlie AM, Atkins DC, Patrick ME, et al. Psychometric analysis and validity of the daily alcohol-related consequences and evaluations measure for young adults. Psychol Assess 2017 Mar;29(3):253-263 [FREE Full text] [doi: 10.1037/pas0000320] [Medline: 27196690]
39. Rose DN, Bond MJ. Identity, stress and substance abuse among young adults. J Subst Use 2009 Jul 12;13(4):268-282. [doi: 10.1080/14659890801912006]
40. Linden-Carmichael AN, van Doren N, Bray BC, Jackson KM, Lanza ST. Stress and affect as daily risk factors for substance use patterns: an application of latent class analysis for daily diary data. Prev Sci 2022 May;23(4):598-607 [FREE Full text] [doi: 10.1007/s11121-021-01305-9] [Medline: 34716891]
41. Terry-McElrath YM, O'Malley PM, Johnston LD. Simultaneous alcohol and marijuana use among U.S. high school seniors from 1976 to 2011: trends, reasons, and situations. Drug Alcohol Depend 2013 Nov 01;133(1):71-79 [FREE Full text] [doi: 10.1016/j.drugalcdep.2013.05.031] [Medline: 23806871]
42. Gunn RL, Aston ER, Sokolovsky AW, White HR, Jackson KM. Complex cannabis use patterns: associations with cannabis consequences and cannabis use disorder symptomatology. Addict Behav 2020 Jun;105:106329 [FREE Full text] [doi: 10.1016/j.addbeh.2020.106329] [Medline: 32044680]
43. Pasch KE, Latimer LA, Cance JD, Moe SG, Lytle LA. Longitudinal bi-directional relationships between sleep and youth substance use. J Youth Adolesc 2012 Sep;41(9):1184-1196 [FREE Full text] [doi: 10.1007/s10964-012-9784-5] [Medline: 22752878]
44. Prince MA, Collins RL, Wilson SD, Vincent PC. A preliminary test of a brief intervention to lessen young adults' cannabis use: episode-level smartphone data highlights the role of protective behavioral strategies and exercise. Exp Clin Psychopharmacol 2020 Apr;28(2):150-156 [FREE Full text] [doi: 10.1037/pha0000301] [Medline: 31144836]
45. Wang D, Wang Y, Wang Y, Li R, Zhou C. Impact of physical exercise on substance use disorders: a meta-analysis. PLoS One 2014 Oct 16;9(10):e110728 [FREE Full text] [doi: 10.1371/journal.pone.0110728] [Medline: 25330437]
46. Kunchay S, Abdullah S. WatchOver: using Apple watches to assess and predict substance co-use in young adults. In: Proceedings of the 2020 ACM International Joint Conference on Pervasive and Ubiquitous Computing and Proceedings of the 2020 ACM International Symposium on Wearable Computers. 2020 Presented at: UbiComp/ISWC '20 Adjunct; September 12-17, 2020; Virtual Event, Mexico. [doi: 10.1145/3410530.3414373]
47. Brooke J. SUS: a 'quick and dirty' usability scale. In: Usability Evaluation In Industry. Boca Raton, FL: CRC Press; 1996.
48. Stoyanov SR, Hides L, Kavanagh DJ, Zelenko O, Tjondronegoro D, Mani M. Mobile app rating scale: a new tool for assessing the quality of health mobile apps. JMIR Mhealth Uhealth 2015 Mar 11;3(1):e27 [FREE Full text] [doi: 10.2196/mhealth.3422] [Medline: 25760773]
49. Yui BH, Jim WT, Chen M, Hsu JM, Liu CY, Lee TT. Evaluation of computerized physician order entry system-a satisfaction survey in Taiwan. J Med Syst 2012 Dec;36(6):3817-3824. [doi: 10.1007/s10916-012-9854-y] [Medline: 22549624]
50. Wright C, Dietze PM, Agius PA, Kuntsche E, Livingston M, Black OC, et al. Mobile phone-based ecological momentary intervention to reduce young adults' alcohol use in the event: a three-armed randomized controlled trial. JMIR Mhealth Uhealth 2018 Jul 20;6(7):e149 [FREE Full text] [doi: 10.2196/mhealth.9324] [Medline: 30030211]
51. Bangor A, Kortum P, Miller J. Determining what individual SUS scores mean: adding an adjective rating scale. J Usability Stud 2009 May 01;4(3):114-123.
52. Liang J, Xian D, Liu X, Fu J, Zhang X, Tang B, et al. Usability study of mainstream wearable fitness devices: feature analysis and system usability scale evaluation. JMIR Mhealth Uhealth 2018 Nov 08;6(11):e11066 [FREE Full text] [doi: 10.2196/11066] [Medline: 30409767]

Abbreviations

EMA: ecological momentary assessment
ICC: intraclass correlation coefficient
mHealth: mobile health
REDCap: Research Electronic Data Capture
SUS: System Usability Scale
uMARS: Mobile Application Rating Scale: User Version

XSL•FO
RenderX

Edited by D Gooch; submitted 20.07.23; peer-reviewed by T Fazzino, C Baxter; comments to author 09.02.24; revised version received 14.03.24; accepted 08.04.24; published 20.06.24.

Please cite as:
Kunchay S, Linden-Carmichael AN, Abdullah S
Using a Smartwatch App to Understand Young Adult Substance Use: Mixed Methods Feasibility Study
JMIR Hum Factors 2024;11:e50795
URL: https://humanfactors.jmir.org/2024/1/e50795
doi:10.2196/50795
PMID:

©Sahiti Kunchay, Ashley N Linden-Carmichael, Saeed Abdullah. Originally published in JMIR Human Factors (https://humanfactors.jmir.org), 20.06.2024. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on https://humanfactors.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Perspectives on the COVID-19 Vaccination Rollout in 17 Countries: Reflexive Thematic and Frequency Analysis Based on the Strengths, Weaknesses, Opportunities, and Threats (SWOT) Framework

Vanja Kopilaš[1*], PhD; Khrystyna Nasadiuk[2*], MD, PhD; Lucia Martinelli[3], PhD; Lenka Lhotska[4,5], PhD; Zoran Todorovic[6,7], MD, PhD; Matjaz Vidmar[8,9], PhD; Helena Machado[10], PhD; Anna Lydia Svalastog[11,12], PhD; Srećko Gajović[13], MD, PhD

[1]Department of Psychology, Faculty of Croatian Studies, University of Zagreb, Zagreb, Croatia
[2]Department of Biochemistry, Danylo Halytsky Lviv National Medical University, Lviv, Ukraine
[3]Science in Society Program, MUSE-Science Museum, Trento, Italy
[4]Czech Institute of Informatics, Robotics and Cybernetics, Czech Technical University in Prague, Prague, Czech Republic
[5]Faculty of Biomedical Engineering, Czech Technical University in Prague, Prague, Czech Republic
[6]University Hospital Medical Center "Bežanijska kosa", Belgrade
[7]Department of Pharmacology, Clinical Pharmacology and Toxicology, Faculty of Medicine, University of Belgrade, Belgrade
[8]Institute for the Study of Science, Technology and Innovation, The University of Edinburgh, Edinburgh, United Kingdom
[9]School of Engineering, The University of Edinburgh, Edinburgh, United Kingdom
[10]Institute for Social Sciences, University of Minho, Braga, Portugal
[11]Østfold University College, Halden, Norway
[12]Centre for Research Ethics and Bioethics, Uppsala University, Uppsala, Sweden
[13]BIMIS-Biomedical Research Center Šalata, University of Zagreb School of Medicine, Zagreb, Croatia
[*]these authors contributed equally

Corresponding Author:
Srećko Gajović, MD, PhD
BIMIS-Biomedical Research Center Šalata
University of Zagreb School of Medicine
Šalata 4
Zagreb, 10000
Croatia
Phone: 385 1 4566 948
Email: srecko.gajovic@hiim.hr

Abstract

Background: As the SARS-CoV-2 virus created a global pandemic and rapidly became an imminent threat to the health and lives of people worldwide, the need for a vaccine and its quick distribution among the population was evident. Due to the urgency, and on the back of international collaboration, vaccines were developed rapidly. However, vaccination rollouts showed different success rates in different countries and some also led to increased vaccine hesitancy.

Objective: The aim of this study was to identify the role of information sharing and context sensitivity in various vaccination programs throughout the initial COVID-19 vaccination rollout in different countries. Moreover, we aimed to identify factors in national vaccination programs related to COVID-19 vaccine hesitancy, safety, and effectiveness. Toward this end, multidisciplinary and multinational opinions from members of the Navigating Knowledge Landscape (NKL) network were analyzed.

Methods: From May to July 2021, 25 completed questionnaires from 27 NKL network members were collected. These contributors were from 17 different countries. The responses reflected the contributors' subjective viewpoints on the status and details of the COVID-19 vaccination rollout in their countries. Contributors were asked to identify strengths, weaknesses, opportunities, and threats (ie, SWOT) of the respective vaccination programs. The responses were analyzed using reflexive thematic analysis, followed by frequency analysis of identified themes according to the represented countries.

XSL•FO
RenderX

Results: The perspectives of NKL network members showed a link between organizational elements of the vaccination rollout and the accompanying societal response, both of which were related to strengths and weaknesses of the process. External sociocultural variables, improved public communication around vaccination-related issues, ethical controversies, and the spread of disinformation were the dominant themes related to opportunities and challenges. In the SWOT 2×2 matrix, *Availability* and *Barriers* emerged as internal categories, whereas *Transparent communication and promotion* and *Societal divide* emerged as key external categories.

Conclusions: Inventory of themes and categories inspired by elements of the SWOT framework provides an informative multidisciplinary perspective for effective implementation of public health strategies in the battle against COVID-19 or any future pandemics of a similar nature.

(JMIR Hum Factors 2024;11:e44258) doi:10.2196/44258

KEYWORDS

SARS-CoV-2 virus; COVID-19 vaccination; pandemic; hesitancy; safety; vaccination; COVID-19; tool; implementation; vaccine hesitancy; effectiveness; sociocultural; communication; disinformation

Introduction

On March 11, 2020, the World Health Organization declared the COVID-19 pandemic [1]. COVID-19 first appeared in December 2019 in Wuhan, China. This disease, caused by the SARS-CoV-2 virus, led to an unprecedented challenge for health institutions that required most countries to integrate their efforts to globally mitigate the spread of the disease [2-4].

Various policies to control the spread of the virus have been adopted in different countries. Some of them were drastic, such as national lockdowns, as well as initiating the widespread use of individual protection devices and means [5]. The individual protective measures included recommending frequent hand washing and application of sanitizers, maintaining social distance, and mandatory wearing of face masks or respirators. However, even a simple measure of covering the face proved to have psychological, cultural, religious, and behavioral implications at both the individual and communal levels [6]. Moreover, the policies aimed to stop the spread of the virus impacted the psychological well-being of the population [7]. These implications should be considered in public campaign strategies aimed at achieving effective public consent toward the adoption of protective measures.

The publication of the genetic sequence of SARS-CoV-2 on January 11, 2020, resulted in the explosion of comprehensive studies on the virus and stirred global research and development activity to develop vaccines against the virus [8]. To accelerate this work, next-generation vaccine technology platforms have been deployed and the first COVID-19 vaccine candidate entered human clinical testing as early as March 16, 2020. In December 2020—in record time and following collaborative efforts of the global scientific community, pharmaceutical companies, and governments—several types and brands of vaccines based on different technologies and mechanisms of action became available for mass deployment [9].

The importance of mass vaccination has been established in the context of previous epidemics such as in the case of smallpox eradication and the incidence reduction of measles and polio [10]. The goal of mass vaccination programs is to interrupt person-to-person disease transmission by surrounding infected people with a high proportion of vaccinated individuals who have developed protective antibodies against the infection (ie, reaching “herd immunity”) [11]. Public health experts have prioritized increased vaccination delivery with the hope to resume socioeconomic activities [12]. According to one study, to reduce the number of confirmed COVID-19 cases and deaths, it was estimated that, on average, the administration of 80 vaccine doses per 100 people was necessary [13]. However, the efficacy of vaccination is challenged by an increasing number of mutated strains, clinically proven possibilities of reinfection, and globally uneven rates of vaccination [14,15].

The vaccination process depends on various societal factors such as vaccine hesitancy, vaccine refusal, practical aspects of its application, and uneven/unequal vaccine rollout [16]. Even prior to the COVID-19 pandemic, the World Health Organization identified vaccine hesitancy as one of the top 10 global threats to public health [17]. After the appearance of COVID-19, this issue has gained a completely different dimension, and several people showed different degrees of vaccination acceptance from total refusal to hesitation, including health workers [18]. Levels of COVID-19 vaccine acceptance and obstacles to its rollout are country- and context-dependent [19]. Research has shown that most people are neither absolutely for nor against COVID-19 vaccines [20,21]. Hence, to begin to understand vaccine acceptance, it is important to gain insight into the reasons behind individual and collective decision-making [22].

SARS-CoV-2 is a novel virus, and various questions about dealing with this threat by mass vaccination emerged during the pandemic, including the efficacy of vaccines, the duration of the vaccine’s effect, and the impact of new virus variants [23]. The rapid vaccine development raised questions regarding safety, availability, and efficacy [24]. This is not surprising given the fact that vaccine development usually takes 10-15 years, whereas COVID-19 vaccines were developed in less than 1 year [25]. In addition, there are various factors that can increase disease spread and mortality rates that seem incoherent with the proposal for a uniform global vaccination rollout. The mortality rates were lower in countries investing more in the health system and vice versa [15,26]. Research from the United States showed that prosperous states with a higher population of older people and a higher number of physicians had a lower rate of vaccine hesitancy compared to that of other states [12].

XSL•FO
RenderX

The availability of vaccines, both in terms of the number of doses and equal distribution, has been an important issue within various countries, involving technical as well as socioeconomic aspects [27]. Timing is also very important since seasonal and environmental factors play important roles in the reduction of COVID-19 symptomatology [15,28]. Due to the numerous factors involved, interdisciplinary collaboration appears to be an appropriate solution to tackle vaccine hesitancy [29].

To facilitate a discussion on a successful vaccination rollout process, in this study, an analysis inspired by the strengths, weaknesses, opportunities, and threats (SWOT) framework was performed to explore perceptions and establish an informative perspective of the vaccination campaigns in 17 countries during the first phase of the mass vaccination programs in the first half of 2021. To facilitate this research, the scholars from the interdisciplinary Navigating Knowledge Landscape (NKL) research network were surveyed between May and July 2021. They were asked to provide information and their own opinions about the vaccination rollout programs in their respective countries. The participating scholars belonged to different disciplines, creating a specific combination of sociological and cultural analytical competences merged with medical and public health expertise. The aim of this interdisciplinary and transnational analysis was to better understand how information-sharing practices and social context were intertwined to coproduce public opinion on vaccination as a response to the COVID-19 pandemic.

SWOT, as a strategic planning framework, is usually used in evaluation of an organization, plan, project, or business activity. It is therefore a significant tool for situation analysis that helps managers identify organizational and environmental factors affecting performance and operations [30]. The framework can be used to identify favorable and unfavorable factors and conditions, solve current problems in a targeted manner, recognize the challenges and obstacles faced, and formulate strategic plans to guide scientific decisions [30-33]. The SWOT framework strives to offer a comprehensive, systematic, and accurate description of the scenario in which a topic is located [34]. SWOT analysis has two dimensions: internal and external. The internal dimension includes organizational factors focusing on strengths and weaknesses, whereas the external dimension includes environmental factors, namely opportunities and threats [30]. Since SWOT analysis is primarily used in organizational studies, our goal was to use its elements as a conceptual and narrative analysis tool where focus was placed on the intertwining viewpoints of social, political, and public health practices. A similar approach has already been applied as a research method in which aspects of the SWOT framework were used to yield more precise and organized data [35]. However, to date, a SWOT-based analysis of the COVID-19 vaccination campaign has only been reported for India and Zimbabwe [36-38]. Therefore, with this study, we aimed to offer a new transdisciplinary and multinational viewpoint of the vaccination process.

Methods

Study Design

The data set included 25 contributions from 27 members of the NKL research network, collected from May to July 2021. These members contributed their viewpoints through a questionnaire aimed at mapping, in a representative manner, the rollout of the vaccination campaigns against SARS-CoV-2 during the early stages when vaccines were available to the general public.

All contributions were collected in a public data set, which is available with open access in Mendeley Data [39].

Study Sample

The 27 contributors were from 17 different countries: Australia, Austria, Croatia, Czech Republic, Germany, Hungary, Italy, Norway, Portugal, Romania, Serbia, Slovenia, South Korea, Sweden, Turkey, Ukraine, and the United Kingdom (including England and Scotland). Three contributions from the same country were received from Slovenia, Sweden, and Portugal; two from Croatia and the United Kingdom; and one contribution from each of the rest of the countries. Two contributions were coauthored (from Australia and Sweden). The contributors come from different academic backgrounds, but most of them are experts in the fields of life sciences, sociology, philosophy, and medicine. However, it is important to note that the contributors were expressing their own opinions and perceptions.

Measures

The questionnaire contained three parts asking about the status and details of COVID-19 vaccination in the respondent's country. Contributors were asked to return a short-text (ie, narrative) answer of 200-300 words per part. In this study, we focused only on the SWOT-related aspect of the responses (ie, Part 1) and the responses to the other parts of the questionnaire (Parts 2 and 3) were considered only to identify the eventual contribution to the SWOT-inspired analysis. SWOT elements were selected among the entire response text during the analysis process. The specific questions are presented in Textbox 1.

XSL•FO
RenderX

Textbox 1. Questionnaire items.

- Part 1: The national vaccination program

Describe the COVID-19 vaccination program in your country: what were its strengths, main weaknesses, opportunities to improve it, and threats to its success?

- Part 2: Public discourse and ethics

How would you describe public responses to your country's vaccination program? What is your impression on the various collective attitudes toward the vaccination program in your country? Were there any ethical issues or concerns around the vaccine program in your country?

- Part 3: Personal experience

What is your personal experience, opinion, or attitude regarding COVID-19 vaccination?

Ethical Approval

Ethical approval for this study was obtained from The University of Edinburgh, Scotland, United Kingdom.

Data Analysis Procedure

To fulfill the study's aims and obtain results, reflexive thematic analysis [40] and descriptive statistics (frequency analysis) of the themes were performed. This method is considered appropriate for exploratory research such as our study. Moreover, flexibility of the thematic analysis and opportunity for theme development seemed a great fit and application for our data set [40,41]. The open-ended questions allowed for formulating responses that enabled the respondents to frame the description of the vaccination process in their countries according to their own personal views.

For the purposes of reflexive thematic analysis, we divided the responses into four categories according to the elements of the SWOT framework. The subcategories of each category were identified and a list of the themes for each SWOT element was established. In a subsequent step, we analyzed the data for patterns and recurring topics. We looked for country-specific differences and similarities in regulations and practices. In addition, close attention was paid to how the experts made sense of their experiences with the vaccination process and how the issues addressed were expressed. In presentation of the research results, focus was placed on themes identified throughout the reflexive thematic analysis. The results were then contextualized based on the existing literature.

Following that, frequency analysis of the identified themes was performed in relation to the corresponding countries. In the case of multiple contributions from the same country referring to the same theme or subtheme, only one data point was counted. The obtained results are presented in the form of tables and graphs.

Results

Thematic Analysis

Overview of Themes

Reflexive thematic analysis of collected contributions was performed independently by two researchers (VK, KN). Through the process of the reflexive thematic analysis [40], the numbers of themes respectively belonging to the elements of strengths, weaknesses, opportunities, and threats were established (Figure 1).

Figure 1. Graphical presentation of the established themes within each of the strengths, weaknesses, opportunities, and threats (SWOT) elements.

XSL•FO
RenderX

Thematic analysis of the vaccination process yielded a nearly even distribution of the four SWOT elements across all included countries and contributors, with 7 themes identified for strengths, 5 themes identified for weaknesses, 6 themes identified for opportunities, and 7 themes identified for threats. In total, analysis of the SWOT elements covered 25 different themes.

The contributors shared their subjective perceptions of the effectiveness of the vaccination campaigns in their countries, which ranged from claims of success to voices of criticism. The United Kingdom was the first country in the world to start the COVID-19 vaccination program in December 2020. Shortly afterward, the vaccine rollout was launched in the United States and the countries of the European Union, albeit with some delay (3 months) in Ukraine. In many countries represented in this study, the vaccination rollout started with some constraints, poor planning/management, and delays with vaccines delivery, but improved over time. In Portugal, an efficient organization of the vaccination process was achieved with the change of the Head of the Vaccination Task Force. In the countries of the European Union, the vaccination process was coordinated with that of other member countries (Croatia), although this collaboration was not always perceived as efficient, as pointed out by a contributor from Sweden.

A successful vaccination program was achieved in the United Kingdom, being respectively described as “overall…a large success” and “an overwhelming success” [39]. The contributions from Portugal and Serbia highly rated the results of the vaccination programs in their respective countries in relation to the high vaccination rate and being ahead of plans/schedule. A relatively successful vaccination process was also reported in Croatia, Hungary, Italy, and Norway. Efficient implementation was noted in Turkey, and an active vaccination process was described in South Korea with major public facilities offering vaccinators discounts or exemptions from paying admission or usage fees. For some other countries, the collected contributions reported low vaccination coverage in the survey period (May to July 2021), including Australia, Romania, and Ukraine. The respondent in Romania specifically reported low coverage for high-risk groups and people over 65 years old. Moreover, very low coverage of rural areas occurred due to lack of local community involvement, especially mayors, policy makers, and family doctors, with some of the latter refusing to dispense vaccines.

Slow rollout of the vaccines was noted in Australia, Austria, and Germany. In Australia, the delayed vaccine rollout has been described as a “vaccine stroll out,” as by July 2021, only 6% of the Australian population had been vaccinated [39]. Moreover, some individuals in priority groups such as older people or those with disabilities living in long-term care homes were still waiting for their second or even their first dose of the vaccine. In addition, in some countries, the vaccination points were hard to access in remote, rural areas (Australia).

If we are to judge vaccination rollout success by looking at the percentage of people who had received at least one dose of the vaccine during the time period corresponding to our data collection, the most successful country in our sample was the United Kingdom, with approximately 70% of the population receiving at least one dose (Multimedia Appendix 1) [42]. The lowest percentage was reported in Ukraine, where only approximately 8% of people had received a single vaccine dose [42].

Strengths

The primary themes related to strengths included (1) societal discussion/consensus on priorities to get the vaccine, (2) defined vaccination strategy/plan, (3) vaccine availability, (4) positive attitudes toward vaccines and the vaccination process, (5) practical aspects of the vaccination solved (eg, medical personnel satisfied, sites easy to access, fast process, no long queues), (6) well-designed public communication campaign on the vaccination process, and (7) flexibility to provide vaccines.

High availability of vaccines was reported in Hungary, Italy, Sweden, and the United Kingdom. Following the controversies around the possible side effects of AstraZeneca vaccines, stocks of the European Union–approved vaccines were excessive in Slovenia. The wide availability of vaccines to whoever wanted them was considered a strength of the vaccination campaigns. In Romania, free vaccination has been offered to everybody who wanted one, including those with Romanian or European citizenship. In Portugal, vaccination was available independent of legal status, including to undocumented migrants. Free vaccination was also offered in Serbia to people from abroad, primarily citizens of neighboring countries, with no restrictions.

Medical workers played a key role in achieving successful vaccination campaigns. Family doctors contributed to the success of the vaccine rollout in Croatia and a helpful approach was reported by the medical staff of the Czech Republic. For Portugal, strong commitment of health care professionals and communication initiatives of the medical doctors to clarify doubts related to the vaccine’s side effects were noted.

Transparent planning and strategies, as well as prioritization of people with higher infection risk or greater vulnerability, were commonly reported strengths of the vaccination programs. In most countries, the prioritization was perceived as fair, although in some countries controversial cases of people from nonpriority groups being vaccinated early also occurred (Portugal, Slovenia). The priority groups in most countries included older adults, those with underlying health conditions, and workers exposed to a high infection risk. By contrast, vaccination of health care professionals has not been prioritized in Sweden. In all countries, the vaccine was provided free of charge, dispensed on a voluntary basis; however, mandatory vaccination was reported for medical workers in Italy and South Korea and for people in high-risk jobs in Australia. Moreover, an easy registration process, owing to easy-to-access platforms such as apps, web pages, or via the phone, was described for Turkey and Ukraine. Automatic enrollment based on medical records via general practitioners (eg, family doctors) was available in the United Kingdom. An efficient registration process in the Czech Republic was also claimed as a strength.

Weaknesses

The primary themes related to weaknesses were as follows: (1) social divide due to the vaccine distribution and side effects,

XSL•FO
RenderX

(2) unclear vaccination strategy/plan, (3) lack of vaccines, (4) negative attitudes or hesitancy toward vaccines and the vaccination process, and (5) barriers to access vaccines.

Lack and shortage of vaccines were emphasized in Ukraine and Turkey, as well as at the beginning of the vaccination rollout in some other countries, where delayed deliveries were also reported. Delayed rollout to the remote Aboriginal communities was noted in Australia. The registration process was essential in achieving an effective rollout of the vaccines. Poor functioning of the distribution organization was highlighted by many participants from different countries, especially in the early stages of vaccination programs. Getting a vaccination appointment was rated as difficult in Sweden.

Trust was pointed out as an important issue in several contributions. A low level of trust in the medical science (Croatia), in the effectiveness and safety of the vaccines (Ukraine, Romania, Slovenia), and in the official authorities (Slovenia) were reported. An overall high level of skepticism in society at large was observed (Germany). Lack of enthusiasm and willingness to be vaccinated or vaccine hesitancy were widely reported (Austria, Norway, Romania, Slovenia, Ukraine). Despite the very successful vaccination process in Serbia, only a small percentage of younger people and health workers were vaccinated in the country. High hesitation among young people was also reported in Slovenia. In contrast, in Ukraine, young people were rather eager to be vaccinated, despite the high level of general hesitancy noted in the country. In Australia, vaccine hesitancy was exacerbated by the risks of side effects reported for the AstraZeneca vaccine.

Lack of clear and coherent communication on how to receive the vaccination was considered an important barrier to access in Slovenia. The need for suitable and unequivocal guidelines about vaccination was stressed in Italy, as constant changes have confused the population and discouraged vaccination, while different rules in different parts of the country and frequent regulation changes were noted to have discouraged vaccination in Germany. Unavailability of scientific information to foreigners/migrants, especially for those not fluent in the local language (eg, the home workers caring for the older population) was stressed for Italy. The digitally based vaccination approach was considered an important barrier to those not having adequate digital abilities. In Sweden, despite having one of the highest internet coverage rates in the world, people living in socially disadvantaged areas, including asylum seekers and migrants, and older adults or people with cognitive impairment who did not master the digital skills required were at risk to be excluded from accessing important information. A low level of digitalization was also mentioned as an obstacle to vaccination success in Romania.

Opportunities

The primary themes related to opportunities were as follows: (1) adding more flexibility; (2) increasing vaccine availability and multiple options for registration; (3) active outreach to marginalized groups, vulnerable citizens, refugees, and ethnic minorities; (4) improving the role of the media (better communication) and national awareness campaigns; (5) information sharing about the usefulness of the vaccination process; and (6) provisions for vaccinated individuals.

The freedom to choose to make an appointment for vaccination, no matter where people were registered (Sweden), and adding more flexibility to accessing vaccination (Croatia) were considered among the opportunities to improve the vaccination rollout.

To motivate people to be vaccinated, financial support (approximately US $30) was offered in Serbia. Vaccination coupons or exemptions from admission or usage fees of public facilities (approximately US $900) were introduced by a National Vaccine Injury Compensation Program in South Korea. In addition, this country also allowed a one-day “vaccination leave” from work to be taken the day after receiving the vaccine, along with an additional one-day leave in the case of experiencing some subsequent side effects [39]. In Ukraine, in the unlikely case that vaccination would cause disability or death, a compensation allowance (approximately US $21,000-27,000) was promised by the government.

Among the opportunities to improve the vaccination process, the freedom to choose among the available vaccine brands/types was recognized as a good strategy to counteract the arising doubts about a certain brand of vaccine (Slovenia, Ukraine). The choice of vaccine brand was also available in Turkey and in Serbia, contributing to successful vaccination campaigns. Finally, a more responsible role of media was mentioned as an opportunity to improve people’s attitude toward vaccination, as pointed out for Croatia and Ukraine. Moreover, in various countries, the respondents suggested that improvement of communication strategies and specific information programs might be crucial to reach vaccine-hesitant citizens and facilitate the vaccine rollout.

In addition to traditional media, social media were noted to play a role. Social media influencers were identified to positively contribute to motivating people to be vaccinated, producing a “crowd effect,” as reported for Croatia and Ukraine, where public figures, such as the President and the health minister gave declarations through the media. Vivid promotions in favor of vaccination by persuasive political and medical discourses, accompanied by enthusiastic argumentations in favor of science and against conspiracy theories and vaccination skepticism, were described for Slovenia.

Threats

The primary themes related to threats were as follows: (1) appearance of new virus strains/lower efficacy of the vaccines, (2) unforeseen side effects of the vaccines, (3) spreading disinformation, (4) ethical controversies, (5) legal controversies, (6) religious controversies, and (7) a change in the behavior of vaccinated individuals that facilitates the spread of infection.

Low trust in the efficacy and safety of the vaccines (Romania, Slovenia, Ukraine); a negative influencing role played by some media communications, especially when stressing the side effects (Serbia, Sweden), and alleged corruption related to the vaccine prioritization (Slovenia) were regarded as relevant threats to be considered for achieving successful mass vaccination campaigns. Insufficient information, disinformation,

XSL•FO
RenderX

or misinformation in the media and on the internet were reported for the Czech Republic, Sweden, and Romania, while development of conspiracy theories about vaccines was pointed out for Slovenia and Ukraine. Disputable communication from the government regarding vaccines and other public health measures such as lockdowns was described for Germany. Lack of adequate public communication strategies was also noted in Slovenia. Failures in communication with people from different cultural groups were reported in Australia.

Ethical concerns associated with the use of leftover doses were pointed out by respondents from Sweden and Portugal, referring to a lack of planning for how to handle leftover vaccines that could not be administered the next day or to the overall mismanagement of vaccine administration. In contrast, the opportunity to get a leftover dose was marked as a strength at the beginning of the vaccination campaign in Ukraine, where this was the only option to be vaccinated for those in nonpriority groups. Confusing messages from religious leaders and local community priests were reported in Romania. Concerns of disobeying the Islamic conduct codes raised by vaccination opponents was described for Turkey, as during the month of Ramadan fears were prompted that vaccination during the fasting period was not acceptable.

Frequency Analysis

Overview

To explore the distribution of the responses by countries, frequency analysis was performed (Figures 2-5). Responses reporting a certain theme are marked in the figures in green color and assigned a value of 1, whereas those that did not mention the theme are marked with light yellow and assigned a value of 0. The total score corresponds to the sum of values of all related responses. Additionally, the average percentage of responses distributed for each element and theme was calculated (Multimedia Appendices 2-5).

Figure 2. Overview of the opinions covering strengths-related themes by country. Green indicates presence of a theme (assigned a value of 1) and yellow indicates absence of the theme (assigned a value of 0). The total score corresponds to the sum of values of all related responses.

	Australia	Austria	Croatia	Czech Republic	Germany	Hungary	Italy	Norway	Portugal	Romania	Serbia	Slovenia	South Korea	Sweden	Turkey	UK-England	UK-Scotland	Ukraine	*Total*
Societal discussion/consensus on priorities to get the vaccine			1							1				1	1	1	1	1	7
Defined vaccination strategy/plan			1				1					1		1	1	1		1	7
Vaccines' availability			1			1					1	1		1		1	1		7
Positive attitudes toward vaccine and vaccination process			1				1		1					1					4
Practical aspects of the vaccination solved										1									1
Well-designed public communication on the vaccination process				1		1					1								3
Flexibility to provide vaccines			1	1					1		1	1			1			1	7
Total	0	0	5	2	0	2	2	0	2	2	3	3	0	4	3	3	2	3	

Figure 3. Overview of the opinions covering weaknesses-related themes by country. Green indicates presence of a theme (assigned a value of 1) and yellow indicates absence of the theme (assigned a value of 0). The total score corresponds to the sum of values of all related responses.

	Australia	Austria	Croatia	Czech Republic	Germany	Hungary	Italy	Norway	Portugal	Romania	Serbia	Slovenia	South Korea	Sweden	Turkey	UK-England	UK-Scotland	Ukraine	*Total*
Social divide due to the vaccine distribution and side effects	1	1	1		1	1	1		1					1					8
Unclear vaccination strategy/plan					1	1			1										3
Lack of vaccines	1		1		1		1		1			1			1			1	8
Negative attitudes or hesitance toward vaccine and vaccination process	1	1				1			1			1		1	1	1		1	9
Barriers to access the vaccine	1		1	1	1	1	1		1	1		1		1	1		1	1	13
Total	4	2	3	1	4	4	3	0	5	1	0	3	0	3	3	1	1	3	

XSL•FO
RenderX

Figure 4. Overview of the opinions covering opportunities-related themes by country. Green indicates presence of a theme (assigned a value of 1) and yellow indicates absence of the theme (assigned a value of 0). The total score corresponds to the sum of values of all related responses.

	Australia	Austria	Croatia	Czech Republic	Germany	Hungary	Italy	Norway	Portugal	Romania	Serbia	Slovenia	South Korea	Sweden	Turkey	UK-England	UK-Scotland	Ukraine	*Total*
Adding more flexibility			■															■	2
Increasing availability of vaccines and multiple options for registration			■		■					■								■	4
Active reach of marginalized groups, vulnerable citizens, refugees, ethnic minorities	■													■					2
Improving role of media (better communication), national awareness campaigns									■	■		■							3
Information spreading about the usefulness of the vaccination process			■						■			■		■				■	5
Provisions for vaccinated individuals											■							■	2
Total	1	0	3	0	1	0	0	0	2	1	1	2	0	2	0	0	0	4	

Figure 5. Overview of the opinions covering threats-related themes by country. Green indicates presence of a theme (assigned a value of 1) and yellow indicates absence of the theme (assigned a value of 0). The total score corresponds to the sum of values of all related responses.

	Australia	Austria	Croatia	Czech Republic	Germany	Hungary	Italy	Norway	Portugal	Romania	Serbia	Slovenia	South Korea	Sweden	Turkey	UK-England	UK-Scotland	Ukraine	*Total*
Appearance of new virus strains/lower efficacy of the vaccine	■										■							■	3
Unforseen side effects of the vaccines											■			■					2
Spreading disinformation	■		■	■						■		■	■	■				■	8
Ethical controversies				■	■		■		■	■		■	■				■	■	9
Legal controversies							■												1
Religious controversies										■									1
Total	2	0	1	2	1	0	2	0	1	3	2	2	2	2	0	0	1	3	

Strengths

Three themes dominated the analysis of strengths, each being covered in 7 reports: societal discussion/consensus on priorities to get the vaccine, defined vaccination strategy/plan, vaccines' availability, and flexibility to provide vaccines.

The strengths theme "societal discussion/consensus on priorities to get the vaccine" was mentioned in Croatia, Romania, Sweden, Turkey, UK Scotland, UK England, and Ukraine. "Defined vaccination strategy/plan" was reported in Croatia, Italy, Slovenia, Sweden, Turkey, UK (England), and Ukraine. "Wide vaccines availability" was indicated in Croatia, Hungary, Serbia, Slovenia, Sweden, and the United Kingdom. "Positive attitudes of the society toward vaccines and vaccination process" were reported for Croatia, Italy, Portugal, and Sweden. Logistic aspects of the vaccination being solved (including satisfaction of the medical personnel, sites easy to access for registration, fast process, no waiting in line) were noted for Romania, while well-designed public communication on the vaccination process was described for the Czech Republic, Hungary, and Serbia. Flexibility to provide vaccines was highlighted as a potential strength in the contributions from Croatia, Czech Republic, Portugal, Serbia, Slovenia, Turkey, and Ukraine. Relatively even distribution was identified across strengths categories with the exception of practical aspects of the vaccination solved that was reported by only one contributor.

Weaknesses

Most frequently reported weaknesses were barriers to access the vaccination (13 reports) and negative attitudes or hesitancy toward vaccines and the vaccination process (9 reports).

Social divide due to the vaccine distribution and side effects were considered weaknesses in Australia, Austria, Croatia, Germany, Hungary, Italy, Portugal, and Sweden. Unclear vaccination strategy/plan was described in Germany, Hungary, and Portugal. Lack of vaccines was noted in Australia, Croatia, Germany, Italy, Portugal, Slovenia, Turkey, and Ukraine (note that the questionnaire addressed these issues only related to the first 6 months of the vaccination campaigns). Negative attitudes or hesitancy toward vaccines and the vaccination process were described in Australia, Austria, Hungary, Portugal, Slovenia, Sweden, Turkey, UK England, and Ukraine, while barriers to access vaccination, including problems with prioritization, registration, and unfair/nontransparent distribution of the vaccines were noted in Australia, Croatia, Germany, Hungary, Italy, Czech Republic, Portugal, Romania, Slovenia, Sweden, Turkey, UK Scotland, and Ukraine. In addition, 72.2% of the contributions reported barriers to access vaccines as a weakness. Conversely, only 16.6% of our sample reported an unclear vaccination strategy/plan as weakness.

XSL•FO
RenderX

Opportunities

The frequencies of the selected opportunities to improve the vaccination process were rather low including a maximum of 5 countries. Adding more flexibility to the vaccination process was mentioned in Croatia and Ukraine; increasing availability of the vaccines and multiple options for registration were mentioned in Croatia, Germany, Romania, and Ukraine; active reach of marginalized groups, vulnerable citizens, refugees, and ethnic minorities was mentioned in Australia and Sweden; improving the role of the media (better communication) and national awareness campaigns were indicated in Portugal, Romania, and Slovenia; information spreading about the usefulness of the vaccination process was highlighted in Croatia, Portugal, Slovenia, Sweden, and Ukraine; and monetary provisions for vaccinated individuals were mentioned in Serbia and Ukraine. Contributors did not report opportunities in large numbers. The highest percentage (27.7%) of responses related to the opportunities-related themes was attributed to spreading information about the usefulness of the vaccination process.

Threats

Concerning the possible threats to a vaccination campaign's success, the contributors from Australia, Serbia, and Ukraine remarked the possible appearance of new virus strains and lower efficacy of the vaccine; unforeseen side effects were noted as possible threats in the contributions from Serbia and Sweden; spreading disinformation were noted or could be concluded from the abstracts from Australia, Croatia, Czech Republic, Romania, Slovenia, South Korea, Sweden, and Ukraine. Other possible threats to vaccination success mentioned were ethical (Czech Republic, Germany, Italy, Portugal, Romania, Slovenia, South Korea, UK Scotland, and Ukraine), along with legal (Italy) and religious controversies (Romania). Ethical controversies (50%) and spreading information (44.4%) were the most highly represented threats-related themes.

In this study, the mainly acknowledged threat feature for achieving successful vaccination campaigns reported by the respondents was related to the likely occurrence of viral mutations, resulting in new virus strains with the ability to escape the immunizing effects of the present available vaccines. This fact has been pointed out as a relevant source of uncertainties and doubts about the vaccines' effectiveness as well as about their overall reliability and utility.

Discussion

Principal Findings

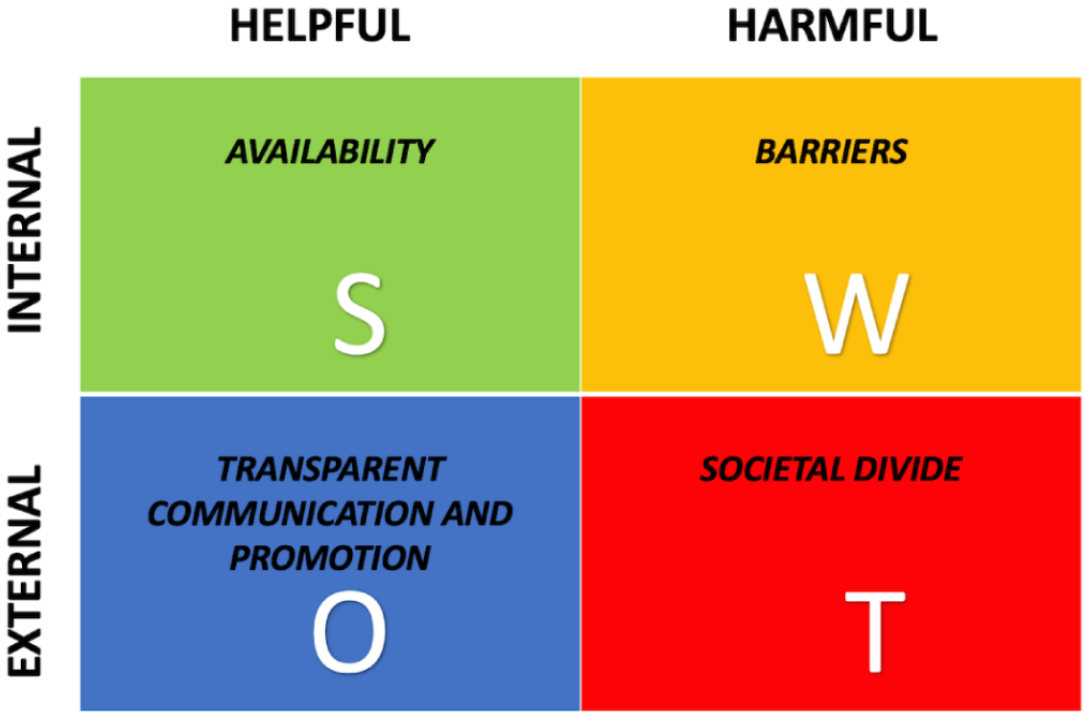
Since its introduction in the 18th century to the present day, vaccination has been one of the most effective tools in the battle against infectious diseases [10,43]. Owing to the high efficacy of vaccines, the public health burden of infectious diseases has been significantly reduced throughout the years [10]. However, despite their proven track record, the phenomenon known as "vaccine hesitancy" has been around almost as long as vaccination itself. This reluctance to accept an injection of an "unknown substance" into the body is exasperated by a need to vaccinate a large number of healthy individuals, including in the case of COVID-19 [44].

This study, based on an analysis of interdisciplinary experts' viewpoints in 17 different countries inspired by the SWOT framework, allowed us to identify 25 themes distributed across the four SWOT elements. To our knowledge, this is the first study to analyze and compare the vaccine rollout process in various different countries. The frequency of the appearance of these themes and their distribution across the countries allowed us to select those that stand out. As the contributions were inspired by the SWOT framework, the presented analysis could be easily synthesized into the four main overarching SWOT categories (Figure 6). With respect to the strengths of the vaccination process, the identified seven themes correspond to a single category referred as *Availability*, being the major strength of the successful vaccination program. When weaknesses were described, the five themes identified could be best described by a single category termed *Barriers*, which were either not recognized or not addressed by the vaccination programs. The external aspects of opportunities described via the six themes identified fit under category *Transparent communication and promotion*, which allows other societal forces to contribute to the vaccination process. Finally, the seven themes describing threats correspond to the *Societal divide* category, where a polarized society has the potential to spoil even well-thought-out initiatives.

We believe that these categories offer the best representation of the most frequently reported themes in each of the SWOT elements. However, due to the intertwining factors present in the vaccination rollout process, it is important to not look at this distribution as a binary (presence/absence) phenomenon. This is particularly relevant when splitting the identified themes into "internal" and "external" categories. In the current SWOT 2×2 matrix (Figure 6), *Availability* and *Barriers* are labeled as internal categories, whereas *Transparent communication and promotion* and *Societal divide* are suggested as external categories [45]. However, within the *Societal divide* category labeled as a threat, there are ethical, religious, and legal controversies reported as important themes. Therefore, one cannot classify a controversy per se as a threat, as controversies can serve just as much as a source of debate with the potential to improve the vaccination process.

XSL•FO
RenderX

Figure 6. Synthesized themes under the strengths (S), weaknesses (W), opportunities (O), and threats (T) framework.



The specific time window when the study was performed corresponds to a relatively early phase of the vaccination process (on average half of the population had been vaccinated in the analyzed countries). This leads to a very specific bias in the submitted data: urgency to tackle an important and pressing issue. We reiterate that our study analyzed the subjective viewpoints of the respondents; hence, some of the themes across these categories were dependent on the various individual, psychological, emotional, and societal aspects specific to the given time window. The sudden appearance of COVID-19 and its rapid spread called for appropriately rapid responses. Considering psychological factors of egocentrism, information availability, social/group confirmation, individual motivation, and emotional affect as foundations of that rapid decision-making process, it is easily possible to misjudge and/or misperceive the key elements of the reasoning arising from the complexity of the situation [46]. However, although there were 27 individual respondents from 17 different countries, our results did not show country-specific differences. Hence, our findings can contribute to the development of strategies that will maximize the promotion of strengths and opportunities while minimizing weaknesses and threats globally.

The identified themes are consistent with the research on this topic [36,38]. The most prominent theme in the existing literature, which was also present in our study, explores the effective medical and public health system measures mapping on the key strengths identified herein. This shows how preparation and prevention strategies work, and how they can be used as a base of the powerful pushback against the spread of COVID-19. Moreover, a positive attitude toward vaccination has been defined as a strength in similar studies in India and Zimbabwe [36,38].

The application of the SWOT framework to complex societal processes can also be seen as a source of confusion. For example, a “strength” is considered as an internal aspect of the process, which can be understood to relate to the vaccination campaign itself. From this perspective, the attitude toward vaccination does not seem to be an internal component but rather an external aspect of SWOT and thus should be more appropriately classified as an opportunity rather than a strength. However, application of the SWOT framework in such complex scenarios requires consideration of the vaccination campaign as part of a sociotechnical system, thus incorporating vital elements of the social environment within the situated practice of vaccination. Combining our findings obtained from individuals from 17 different countries with previous research, it can be concluded that good organization that addresses the availability of vaccines coupled with an engaging societal discussion would represent a key strength/opportunity of the vaccination process.

A lack/shortage of vaccines combined with various logistical challenges have been reported as major issues for the success of vaccination campaigns within previous research [47,48]. The demand-supply gap combined with lack of knowledge and supporting infrastructures have been reported as particular weaknesses [36,38]. Compounding unequal vaccine distribution with unknown disease progression and an uncertain response to the vaccine seems to be the biggest barrier in the vaccine rollout [13]. Similarly, the respondents of this study recognized the practical issues of availability and fairness of distribution, and coupled these issues with the related attitudes and social division. This points to the fact that social distrust needs to be addressed within a vaccination plan as a major barrier. For both strengths and weaknesses, no clear geographical divide was present.

Increasing the public awareness about the vaccine effects through transparent communication and promotion stood out as a key opportunity-related theme. Communication reports on

XSL•FO
RenderX

the widespread acceptance of COVID-19 vaccines have shown to be effective tools to further increase vaccine acceptance [49]. Moreover, in an attempt to promote vaccination, some public figures have been vaccinated on television [49]. It is interesting to see that people who used mainstream media outlets as their major source of information on health were more likely to get vaccinated [50]. Our data support the notion that transparent information-sharing about biological mechanisms, efficacy, as well as side effects of the vaccines motivates people to join vaccination programs. Previous research has identified the potentially influential role of media in increasing people's trust in vaccines when they hear politicians, celebrities, or other famous people talking positively about them [36]. Trust in vaccines, medical science, and medical professionals—together with other involved stakeholders, including government and policy makers—was highlighted in the analyzed contributions as an important factor. These findings align with previous research that found lack of communication from trusted providers and community leaders as one of the main reasons for low COVID-19 vaccination rates [44]. Communication of vaccine information and promotion of its uptake in the digital era includes the use of social networks [51]. However, the use of social networks is also associated with risks due to the wildfire-like dynamics of rumors in the digital environment and issues with unknown algorithms used by for-profit entities filtering information [52,53]. Social networks are expected to drive healthy public debates; however, they instead frequently reinforce like-minded "bubbles" and increase polarization [53,54].

Discussions on matters of autonomy (an individual's right to choose) and state power have always been at the center of public health ethical dilemmas [29]. In the specific case of COVID-19 vaccines, besides the tensions between public health and individual interest/autonomy, other ethical challenges relate to the rapid design and testing of vaccines and who gets the vaccines (first) [55]. Public health authorities need to implement efficient, flexible, responsive, and resilient strategies to successfully fight the pandemic and raise awareness of all of the dangers arising with this disease [56]. Surprisingly, in our findings, the question of one's autonomy did not crystalize as a theme. Instead, other ethical controversies and spreading of disinformation were found to be the most frequently reported themes within the threats element. In the present digital era, information accessibility is at its peak; however, it is important be aware of the source of the information given that rumors and fake news are rampant [17,57].

When discussing the threats element in the SWOT framework, it was interesting that unforeseen side effects of the vaccine have not been considered as the most prominent threat theme, whereas other research shows that the most common reason for vaccine hesitancy or refusal is due to the concerns related to the side effects/safety [50,58,59]. The emergence of new virus strains was mentioned as a threat, since they decrease the efficacy of the vaccine and hence can contribute to the further spread of COVID-19. As people were already worried about the lack of information about safety, testing, and efficacy of COVID-19 vaccines, the new variants were seen as a contributor to the negative perception of the vaccines in society [15,60]. The synthesizing category for the threat element of *Societal divide* implies that social polarizations have the potential to paralyze a society when facing a complex public health crisis. Here, it should be stated that silencing the controversies is certainly not the path to avoid such an outcome. A society where controversies are not openly discussed is not without these controversies, but rather this situation would give rise to potentially dangerous and isolation subcultures. Although *Societal divide* was recognized as a threat in our sample, there were no clear examples where this has significantly directly influenced the vaccination process. Consequently, although the awareness of controversies as a potential threat was voiced, if the social environment is developed within the context of *Transparent communication and promotion* (opportunity), the *Social divide* may never reach the level of polarization to create adverse effects on public health campaigns.

Study Limitations

The collected responses represent the subjective viewpoints of experts who volunteered to take part in the study. Therefore, extrapolation to the national level must be drawn out with caution. In addition, due to the lack of research using this same methodology and implementing it on a multinational level, there were no relevant studies to make direct comparisons with and contrast conclusions. Moreover, SWOT analysis was not performed in its original form addressing organizational dynamics. Instead, this thematic analysis of expert viewpoints was only inspired by the SWOT framework. Therefore, the results of this study should be further examined and more research is needed on this topic in general. Further studies could consider interdisciplinary and multinational frameworks to find the best practice in public health policies that could yield improved vaccination rollout results globally.

Conclusion

This study was based on a collection of short responses to a specifically designed questionnaire, written by researchers from different countries and fields of expertise, thus bringing together multidisciplinary and cross-national opinions on vaccination rollout. This represents the first analysis of the vaccination process in 17 different countries inspired by the SWOT framework. The obtained results highlight the connection between organizational aspects of the vaccination rollout and corresponding societal response, both being related to the strengths and weaknesses of the process. The opportunities and threats corresponded to external societal factors, better public communication of vaccination-related issues, ethical controversies, and the spread of disinformation. The inventory of 25 SWOT-related themes and the resulting 2×2 SWOT matrix represents an approximate best-practice viewpoint for the successful implementation of public health policies—as represented by this multidisciplinary team—in the fight against COVID-19.

XSL•FO
RenderX

Acknowledgments

This study represents an activity of the interdisciplinary network Navigating Knowledge Landscapes. The study was supported by the European Union through the European Regional Development Fund, as the Scientific Centre of Excellence for Basic, Clinical and Translational Neuroscience, under grant agreement KK.01.1.1.07.0071, project "Synergy of molecular markers and multimodal in vivo imaging during preclinical assessment of the consequences of the ischemic stroke (SineMozak)" (to SG).

Data Availability

The data sets supporting the results presented in this study can be found in the online repository [39].

Authors' Contributions

VK, KN, MV, and SG designed the study. VK, KN, and LM performed data acquisition, organization and analysis, and wrote the first version of the manuscript. VK, KN, LM, LL, ZT, MV, HM, ALS, and SG contributed to the interpretation of the data collected, framed the results, and critically revised the manuscript. All authors approved the submission to the journal.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Share of people from countries in our sample who received at least one dose of COVID-19 vaccine (taken from Our World in Data. Coronavirus (COVID-19) Vaccinations. 2021 [42]).
[PNG File , 438 KB - humanfactors_v11i1e44258_app1.png]

Multimedia Appendix 2
Average percentage of opinions covering strengths-related themes.
[PNG File , 151 KB - humanfactors_v11i1e44258_app2.png]

Multimedia Appendix 3
Average percentage of opinions covering weaknesses-related themes.
[PNG File , 126 KB - humanfactors_v11i1e44258_app3.png]

Multimedia Appendix 4
Average percentage of opinions covering opportunities-related themes.
[PNG File , 140 KB - humanfactors_v11i1e44258_app4.png]

Multimedia Appendix 5
Average percentage of opinions covering threats-related themes.
[PNG File , 134 KB - humanfactors_v11i1e44258_app5.png]

References

1. Archived: WHO Timeline - COVID-19. World Health Organization. 2020 Apr. URL: https://www.who.int/news/item/27-04-2020-who-timeline---covid-19 [accessed 2023-07-03]
2. Zhou P, Yang X, Wang X, Hu B, Zhang L, Zhang W, et al. A pneumonia outbreak associated with a new coronavirus of probable bat origin. Nature 2020 Mar;579(7798):270-273 [FREE Full text] [doi: 10.1038/s41586-020-2012-7] [Medline: 32015507]
3. Usher K, Durkin J, Bhullar N. The COVID-19 pandemic and mental health impacts. Int J Ment Health Nurs 2020 Jun;29(3):315-318 [FREE Full text] [doi: 10.1111/inm.12726] [Medline: 32277578]
4. Satici B, Saricali M, Satici SA, Griffiths MD. Intolerance of uncertainty and mental wellbeing: serial mediation by rumination and fear of COVID-19. Int J Ment Health Addict 2022;20(5):2731-2742 [FREE Full text] [doi: 10.1007/s11469-020-00305-0] [Medline: 32427165]
5. Indirili M, Di Maio V, Martinelli L. An analysis of Italian resilience during COVID-19 pandemic: first phase from January to June 2020. In: De Pascale F, editor. Geographies of the Anthropocene. Lago, Italy: Il Sileno Edizione; 2021.
6. Martinelli L, Kopilaš V, Vidmar M, Heavin C, Machado H, Todorović Z, et al. Face masks during the COVID-19 pandemic: a simple protection tool with many meanings. Front Public Health 2020 Jan 13;8:606635 [FREE Full text] [doi: 10.3389/fpubh.2020.606635] [Medline: 33520918]
7. Kopilaš V, Hasratian AM, Martinelli L, Ivkić G, Brajković L, Gajović S. Self-perceived mental health status, digital activity, and physical distancing in the context of lockdown versus not-in-lockdown measures in Italy and Croatia: cross-sectional

XSL•FO
RenderX

study in the early ascending phase of the COVID-19 pandemic in March 2020. Front Psychol 2021 Feb 4;12:621633 [FREE Full text] [doi: 10.3389/fpsyg.2021.621633] [Medline: 33613398]
8. Forster P, Forster L, Renfrew C, Forster M. Phylogenetic network analysis of SARS-CoV-2 genomes. Proc Natl Acad Sci U S A 2020 Apr 28;117(17):9241-9243 [FREE Full text] [doi: 10.1073/pnas.2004999117] [Medline: 32269081]
9. Thanh Le T, Andreadakis Z, Kumar A, Gómez Román R, Tollefsen S, Saville M, et al. The COVID-19 vaccine development landscape. Nat Rev Drug Discov 2020 May;19(5):305-306. [doi: 10.1038/d41573-020-00073-5] [Medline: 32273591]
10. Saleh A, Qamar S, Tekin A, Singh R, Kashyap R. Vaccine development throughout history. Cureus 2021 Jul;13(7):e16635 [FREE Full text] [doi: 10.7759/cureus.16635] [Medline: 34462676]
11. Randolph HE, Barreiro LB. Herd immunity: understanding COVID-19. Immunity 2020 May 19;52(5):737-741 [FREE Full text] [doi: 10.1016/j.immuni.2020.04.012] [Medline: 32433946]
12. Goel RK, Jones JR, Saunoris JW. Explaining vaccine hesitancy: a COVID-19 study of the United States. MDE Manage Decis Econ 2023;44(2):1073-1087 [FREE Full text] [doi: 10.1002/mde.3732] [Medline: 36247212]
13. Coccia M. Optimal levels of vaccination to reduce COVID-19 infected individuals and deaths: a global analysis. Environ Res 2022 Mar;204(Pt C):112314 [FREE Full text] [doi: 10.1016/j.envres.2021.112314] [Medline: 34736923]
14. Patel V, Shirbhate E, Rajak H. Coronavirus reinfections: an outlook on evidences and effects. In: Kaklauskas A, Abraham A, Okoye K, Guggari S, editors. Lessons from COVID-19: impact on healthcare systems and technology. Amsterdam: Elsevier Academic Press; 2022:19-40.
15. Coccia M. COVID-19 vaccination is not a sufficient public policy to face crisis management of next pandemic threats. Public Organiz Rev 2022 Oct 17;23(4):1353-1367. [doi: 10.1007/s11115-022-00661-6]
16. Leigh JP, Moss SJ, White TM, Picchio CA, Rabin KH, Ratzan SC, et al. Factors affecting COVID-19 vaccine hesitancy among healthcare providers in 23 countries. Vaccine 2022 Jul 29;40(31):4081-4089 [FREE Full text] [doi: 10.1016/j.vaccine.2022.04.097] [Medline: 35654620]
17. Teitler-Regev S, Hon-Snir S. COVID-19 vaccine hesitancy in Israel immediately before the vaccine operation. Yale J Biol Med 2022 Jun;95(2):199-205 [FREE Full text] [Medline: 35782475]
18. Gur-Arie R, Jamrozik E, Kingori P. No jab, no job? Ethical issues in mandatory COVID-19 vaccination of healthcare personnel. BMJ Glob Health 2021 Feb 17;6(2):e004877 [FREE Full text] [doi: 10.1136/bmjgh-2020-004877] [Medline: 33597280]
19. de Figueiredo A, Simas C, Larson HJ. COVID-19 vaccine acceptance and its socio-demographic and emotional determinants: a multi-country cross-sectional study. Vaccine 2023 Jan 09;41(2):354-364 [FREE Full text] [doi: 10.1016/j.vaccine.2022.10.051] [Medline: 36414475]
20. Schwarzinger M, Watson V, Arwidson P, Alla F, Luchini S. COVID-19 vaccine hesitancy in a representative working-age population in France: a survey experiment based on vaccine characteristics. Lancet Public Health 2021 Apr;6(4):e210-e221 [FREE Full text] [doi: 10.1016/S2468-2667(21)00012-8] [Medline: 33556325]
21. Verger P, Peretti-Watel P. Understanding the determinants of acceptance of COVID-19 vaccines: a challenge in a fast-moving situation. Lancet Public Health 2021 Apr;6(4):e195-e196 [FREE Full text] [doi: 10.1016/S2468-2667(21)00029-3] [Medline: 33556329]
22. Machingaidze S, Wiysonge CS. Understanding COVID-19 vaccine hesitancy. Nat Med 2021 Aug;27(8):1338-1339. [doi: 10.1038/s41591-021-01459-7] [Medline: 34272500]
23. Forni G, Mantovani A, COVID-19 Commission of Accademia Nazionale dei Lincei, Rome. COVID-19 vaccines: where we stand and challenges ahead. Cell Death Differ 2021 Feb;28(2):626-639 [FREE Full text] [doi: 10.1038/s41418-020-00720-9] [Medline: 33479399]
24. Rosenthal S, Cummings CL. Influence of rapid COVID-19 vaccine development on vaccine hesitancy. Vaccine 2021 Dec 20;39(52):7625-7632 [FREE Full text] [doi: 10.1016/j.vaccine.2021.11.014] [Medline: 34802786]
25. Kashte S, Gulbake A, El-Amin Iii SF, Gupta A. COVID-19 vaccines: rapid development, implications, challenges and future prospects. Hum Cell 2021 May 07;34(3):711-733 [FREE Full text] [doi: 10.1007/s13577-021-00512-4] [Medline: 33677814]
26. Coccia M. High health expenditures and low exposure of population to air pollution as critical factors that can reduce fatality rate in COVID-19 pandemic crisis: a global analysis. Environ Res 2021 Aug;199:111339 [FREE Full text] [doi: 10.1016/j.envres.2021.111339] [Medline: 34029545]
27. Mathieu E, Ritchie H, Ortiz-Ospina E, Roser M, Hasell J, Appel C, et al. A global database of COVID-19 vaccinations. Nat Hum Behav 2021 Jul;5(7):947-953. [doi: 10.1038/s41562-021-01122-8] [Medline: 33972767]
28. Coccia M. COVID-19 pandemic over 2020 (withlockdowns) and 2021 (with vaccinations): similar effects for seasonality and environmental factors. Environ Res 2022 May 15;208:112711 [FREE Full text] [doi: 10.1016/j.envres.2022.112711] [Medline: 35033552]
29. Harrison EA, Wu JW. Vaccine confidence in the time of COVID-19. Eur J Epidemiol 2020 Apr;35(4):325-330 [FREE Full text] [doi: 10.1007/s10654-020-00634-3] [Medline: 32318915]
30. Pickton DW, Wright S. What's SWOT in strategic analysis? Strat Change 1998 Mar;7(2):101-109. [doi: 10.1002/(sici)1099-1697(199803/04)7:2<101::aid-jsc332>3.0.co;2-6]

XSL•FO
RenderX

31. Wang J, Wang Z. Strengths, Weaknesses, Opportunities and Threats (SWOT) analysis of China's prevention and control strategy for the COVID-19 epidemic. Int J Environ Res Public Health 2020 Mar 26;17(7):2235 [FREE Full text] [doi: 10.3390/ijerph17072235] [Medline: 32225019]
32. Zhou Y, Bai L, Guo H, Guo S, Han X, Yue NJ, et al. SWOT analysis and preliminary study on prevention and control management of temporary integrated isolation ward during COVID-19 outbreak. Front Public Health 2021 Mar 15;9:558565 [FREE Full text] [doi: 10.3389/fpubh.2021.558565] [Medline: 33791264]
33. Miranda AE, Rosadas C, Assone T, Pereira GFM, Vallinoto ACR, Ishak R. Strengths, Weaknesses, Opportunities and Threats (SWOT) analysis of the implementation of public health policies on HTLV-1 in Brazil. Front Med 2022 Apr 7;9:859115 [FREE Full text] [doi: 10.3389/fmed.2022.859115] [Medline: 35462992]
34. Jasiulewicz-Kaczmarek M. SWOT analysis for Planned Maintenance strategy-a case study. IFAC-PapersOnLine 2016;49(12):674-679. [doi: 10.1016/j.ifacol.2016.07.788]
35. Asgary NH, McNulty RE. Contributions and challenges in the struggle to end poverty: the case of Kiva. Inf Technol Develop 2017 Apr 06;23(2):367-387. [doi: 10.1080/02681102.2016.1233856]
36. Sharma P, Pardeshi G. Rollout of COVID-19 vaccination in India: a SWOT analysis. Disaster Med Public Health Prep 2022;16(6):2310-2313 [FREE Full text] [doi: 10.1017/dmp.2021.111] [Medline: 33820581]
37. Thakur H. A strengths, weaknesses, opportunities, and threats analysis of public health in India in the context of COVID-19 pandemic. Indian J Community Med 2021;46(1):1. [doi: 10.4103/ijcm.ijcm_1053_20]
38. Murewanhema G, Musuka G, Denhere K, Chingombe I, Mapingure MP, Dzinamarira T. The landscape of COVID-19 vaccination in Zimbabwe: a narrative review and analysis of the strengths, weaknesses, opportunities and threats of the programme. Vaccines 2022 Feb 09;10(2):262 [FREE Full text] [doi: 10.3390/vaccines10020262] [Medline: 35214720]
39. Kopilaš V, Buzas N, Carmesund U, Ceausu I, Chenhall R, de FC, et al. The collection of narratives on strategies, ethics and public response of COVID-19 national vaccination process written by members of the scholarly association Navigating Knowledge Landscapes Network. Mendeley Data. 2022. URL: https://data.mendeley.com/datasets/bwdf6v5wgr/1 [accessed 2022-09-06]
40. Braun V, Clarke V. Using thematic analysis in psychology. Qual Res Psychol 2006 Jan;3(2):77-101. [doi: 10.1191/1478088706qp063oa]
41. Byrne D. A worked example of Braun and Clarke's approach to reflexive thematic analysis. Qual Quant 2021 Jun 26;56(3):1391-1412. [doi: 10.1007/s11135-021-01182-y]
42. Coronavirus (COVID-19) Vaccinations. Our World in Data. URL: https://ourworldindata.org/covid-vaccinations [accessed 2023-07-03]
43. Plotkin S. History of vaccination. Proc Natl Acad Sci U S A 2014 Aug 26;111(34):12283-12287 [FREE Full text] [doi: 10.1073/pnas.1400472111] [Medline: 25136134]
44. Razai MS, Chaudhry UAR, Doerholt K, Bauld L, Majeed A. Covid-19 vaccination hesitancy. BMJ 2021 May 20;373:n1138. [doi: 10.1136/bmj.n1138] [Medline: 34016653]
45. Moltu C, Stefansen J, Svisdahl M, Veseth M. Negotiating the coresearcher mandate - service users' experiences of doing collaborative research on mental health. Disabil Rehabil 2012;34(19):1608-1616. [doi: 10.3109/09638288.2012.656792] [Medline: 22489612]
46. Blattman C. Why we fight: the roots of war and the paths to peace. Westminister, London: Penguin Books; 2023.
47. Torjesen I. Covid-19 vaccine shortages: what is the cause and what are the implications? BMJ 2021 Mar 19;372:n781. [doi: 10.1136/bmj.n781] [Medline: 33741547]
48. Liu K, Lou Y. Optimizing COVID-19 vaccination programs during vaccine shortages. Infect Dis Model 2022 Mar;7(1):286-298 [FREE Full text] [doi: 10.1016/j.idm.2022.02.002] [Medline: 35233475]
49. Moehring A, Collis A, Garimella K, Rahimian MA, Aral S, Eckles D. Providing normative information increases intentions to accept a COVID-19 vaccine. Nat Commun 2023 Jan 09;14(1):126. [doi: 10.1038/s41467-022-35052-4] [Medline: 36624092]
50. Viswanath K, Bekalu M, Dhawan D, Pinnamaneni R, Lang J, McLoud R. Individual and social determinants of COVID-19 vaccine uptake. BMC Public Health 2021 Apr 28;21(1):818 [FREE Full text] [doi: 10.1186/s12889-021-10862-1] [Medline: 33910558]
51. Cascini F, Pantovic A, Al-Ajlouni YA, Failla G, Puleo V, Melnyk A, et al. Social media and attitudes towards a COVID-19 vaccination: a systematic review of the literature. EClinicalMedicine 2022 Jun;48:101454 [FREE Full text] [doi: 10.1016/j.eclinm.2022.101454] [Medline: 35611343]
52. Gottlieb M, Dyer S. Information and disinformation: social media in the COVID-19 crisis. Acad Emerg Med 2020 Jul;27(7):640-641 [FREE Full text] [doi: 10.1111/acem.14036] [Medline: 32474977]
53. Kopilaš V, Gajović S. Wildfire-like effect of a WhatsApp campaign to mobilize a group of predominantly health professionals with a university degree on a health issue: infodemiology study. J Med Internet Res 2020 Aug 10;22(8):e17051 [FREE Full text] [doi: 10.2196/17051] [Medline: 32442138]
54. Xie L, Wang D, Ma F. Analysis of individual characteristics influencing user polarization in COVID-19 vaccine hesitancy. Comput Human Behav 2023 Jun;143:107649 [FREE Full text] [doi: 10.1016/j.chb.2022.107649] [Medline: 36683861]

XSL•FO
RenderX

55. Raus K, Mortier E, Eeckloo K. Ethical issues in managing the COVID-19 pandemic. Bioethics 2021 Jul 05;35(6):581-588 [FREE Full text] [doi: 10.1111/bioe.12877] [Medline: 33951206]
56. Coccia M. Pandemic prevention: lessons from COVID-19. Encyclopedia 2021 May 31;1(2):433-444. [doi: 10.3390/encyclopedia1020036]
57. Roy DN, Azam MS, Biswas M, Islam E. Potential factors influencing COVID-19 vaccine acceptance and hesitancy among university students in Bangladsh: a cross-sectional comparative study. Epidemiol Infect 2022 Dec 20;151:e11 [FREE Full text] [doi: 10.1017/S0950268822001820] [Medline: 36537315]
58. Solís Arce JS, Warren SS, Meriggi NF, Scacco A, McMurry N, Voors M, et al. COVID-19 vaccine acceptance and hesitancy in low- and middle-income countries. Nat Med 2021 Aug;27(8):1385-1394 [FREE Full text] [doi: 10.1038/s41591-021-01454-y] [Medline: 34272499]
59. Hong L, Jin Z, Xu K, Shen G, Zou Y, Li R, et al. COVID-19 vaccine uptake and vaccine hesitancy in rural-to-urban migrant workers at the first round of COVID-19 vaccination in China. BMC Public Health 2023 Jan 20;23(1):139 [FREE Full text] [doi: 10.1186/s12889-023-15068-1] [Medline: 36658507]
60. Abou-Arraj NE, Maddah D, Buhamdan V, Abbas R, Jawad NK, Karaki F, et al. Perceptions of, and obstacles to, SARS-CoV-2 vaccination among adults in Lebanon: cross-sectional online survey. JMIR Form Res 2022 Dec 14;6(12):e36827 [FREE Full text] [doi: 10.2196/36827] [Medline: 36383635]

Abbreviations

NKL: Navigating Knowledge Landscape
SWOT: strengths, weaknesses, opportunities, and threats

Edited by A Kushniruk; submitted 12.11.22; peer-reviewed by M Coccia, B Hausman, Y Mejova; comments to author 01.06.23; revised version received 16.07.23; accepted 26.09.23; published 19.02.24.

Please cite as:
Kopilaš V, Nasadiuk K, Martinelli L, Lhotska L, Todorovic Z, Vidmar M, Machado H, Svalastog AL, Gajović S
Perspectives on the COVID-19 Vaccination Rollout in 17 Countries: Reflexive Thematic and Frequency Analysis Based on the Strengths, Weaknesses, Opportunities, and Threats (SWOT) Framework
JMIR Hum Factors 2024;11:e44258
URL: https://humanfactors.jmir.org/2024/1/e44258
doi:10.2196/44258
PMID:38373020

©Vanja Kopilaš, Khrystyna Nasadiuk, Lucia Martinelli, Lenka Lhotska, Zoran Todorovic, Matjaz Vidmar, Helena Machado, Anna Lydia Svalastog, Srećko Gajović. Originally published in JMIR Human Factors (https://humanfactors.jmir.org), 19.02.2024. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on https://humanfactors.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

The Role of Social Media in the Experiences of COVID-19 Among Long-Hauler Women: Qualitative Study

Camryn Garrett[1], MPH; Atefeh Aghaei[1], MA, PhD; Abhishek Aggarwal[1], BA; Shan Qiao[1], MA, PhD

Department of Health Promotion, Education, and Behavior, Arnold School of Public Health, University of South Carolina, Columbia, SC, United States

Corresponding Author:
Camryn Garrett, MPH
Department of Health Promotion, Education, and Behavior
Arnold School of Public Health
University of South Carolina
915 Greene Street
Columbia, SC, 29208
United States
Phone: 1 803 777 6844
Email: camrynmg@email.sc.edu

Abstract

Background: The extant literature suggests that women are more vulnerable to COVID-19 infection and at higher risk for developing long COVID. Due to pandemic mitigation recommendations, social media was relied upon for various aspects of daily life, likely with differences of usage between genders.

Objective: This study aimed to explore the role and functions of social media in the lives of long-hauler women.

Methods: Participants were purposively snowball-sampled from an online health promotion intervention for long-hauler women with COVID-19 from March to June 2021. During this time, one-on-one, semistructured interviews were conducted online until data saturation was agreed to have been achieved (ie, 15 interviews). Interview transcripts and field notes were analyzed using an emergent, inductive approach.

Results: In total, 15 women were enrolled. The main roles of social media included facilitating support group participation, experience sharing, interpersonal connections, and media consumption. Emergent themes demonstrated that participants rely on social media to fulfill needs of emotional support, social engagement, spirituality, health planning, information gathering, professional support, and recreationally for relaxation. As long-hauler women turn to social media to discuss symptom and health management as well as the intention to vaccinate, this study demonstrates both the associated benefits (ie, decreased isolation) and challenges (ie, misinformation, rumination, resentment, jealousy).

Conclusions: The public health implications of these findings support the development of gender-tailored health promotion interventions that leverage the benefits of social media, while mitigating the negative impacts, for women with long COVID.

(JMIR Hum Factors 2024;11:e50443) doi:10.2196/50443

KEYWORDS

COVID-19; long COVID; long-haulers; women; gender; social media; digital media; qualitative study

Introduction

The COVID-19 Context

The COVID-19 pandemic, caused by SARS-CoV-2, was declared on March 11, 2020, and was accompanied by recommendations to implement prevention measures, such as masking, vaccination, testing, social distancing, isolation, and quarantine [1-5]. In the United States, as of June 16, 2023, there had been more than 103.4 million cases of COVID-19, of which more than 6.1 million required hospitalization and more than 1.1 million resulted in death [6,7]. Of those who become infected, about 30% develop postacute sequela SARS-CoV-2 (PASC), also known as chronic COVID-19 or long COVID, characterized by symptoms of varying severity that persist for 4 weeks or more after infection (eg, chronic fatigue, pain, cognitive dysfunction, muscle deconditioning, impaired concentration, and persistent ageusia and anosmia); these patients are commonly referred to as long-haulers [8-13]. Overall, women have been found to be more likely than men to develop long COVID (ie, 9.4% vs 5.5%) [14-16]. This

XSL•FO
RenderX

disproportionate trend requires further investigation into the differential experiences of long COVID among women.

Long COVID Among Women

The literature shows that there may be an association between biological sex and COVID-19 infection and recovery; however, this fails to consider the role of gender, the social environment, and gendered social norms [17]. For instance, women primarily constitute social assistance and health care workforces and face increased expectations of caregiving in the family setting, increasing the risk of COVID-19 infection [18]. As long COVID results from COVID-19 infection, differential exposure and incidence among women predisposes them to the risk for developing persisting symptoms [19]. Persistent symptoms experienced more so by long-hauler women include fatigue, difficulty breathing, muscle pain, and cognitive dysfunction, as well as the negative psychosocial outcomes of anxiety, depression, and posttraumatic stress disorder (PTSD), particularly among those who have been hospitalized [20,21].

Social Media and COVID-19

Due to stay-at-home orders and the prioritization of social distancing as the primary means to prevent the spread of COVID-19, in the online environment, social media emerged as a key tool to adjust to the new normal, necessitating research on its roles and functions. Overall, in the United States, 97% of Americans indicate owning a cell phone of any kind, 85% indicate owning a smartphone, and 85% of US households have a broadband internet connection [22,23]. Using these technologies, 85% of US adults indicate going online at least once a day and 31% indicate that they are online “almost constantly” [22]. Among mixed findings in the COVID-19 literature, women report higher usage of social media compared to men, with assumedly differential motivations for engagement and use of platform functions [24]. Due to the prevalence of individuals being online for work and personal use, it is necessary to evaluate the role of technology and social media within the context of the pandemic and, specifically, the experiences of long-hauler women [25].

Among long-haulers specifically, social media played a vital role in developing the long-hauler identity and encouraging clinical acknowledgement. The term “long COVID” originates from social media users’ posts online [26]. Posts sharing long COVID experiences typically include a diagnosis or test result, the symptoms experienced, the length of time symptoms have persisted, an emotional response, and information and resources [27]. The growing conversations among long-haulers on social media shifted the experience of long COVID from anecdotal, exposing an invisible disability, to clinical [26,27]. With the creation of a shared identity, long-haulers were able to identify one another and further subdivided themselves into categories accounting for their intersecting identities (eg, long-hauler, woman, and mother).

Social media has been used to mitigate the impacts of lost social connections, social distancing, and isolation [21,22]. In the literature, social media has been demonstrated as a key tool used to maintain social connections, while adhering to social distancing recommendations, limiting feelings of isolation [28]. Additionally, as loneliness has been associated with decreased use of healthful coping behaviors, social media has been found to mediate the association during periods of isolation, such as during the COVID-19 pandemic [29]. There are linkages between daily use of social media and lowered measures of social isolation, as well as inversely with infrequent use of social media and higher measures of isolation [30]. Despite the positive outcomes associated with using social media during the pandemic, downfalls remain. For example, in a study focusing on older adults, internet use reflected coping efforts but did not necessarily enhance or sufficiently improve well-being [31]. Due to the mixed effects and roles of social media, throughout the pandemic, there is a unique opportunity for researchers to investigate the role of social media in social connections, isolation, and support, as well as in perpetuating access to information or misinformation among long-hauler women [32].

Overall, the literature suggests that social media sites impact users’ ability to maintain social connections, seek social support, and access information, as well as affect isolation, social comparison, and the spread of misinformation [33,34]. According to the information systems literature, gender is associated with differential motivations to use social media sites (eg, relational uses for women vs information gathering for men) and differential perceptions of information shared online [35,36]. To the best of our knowledge, despite the gendered associations relevant to social media use in other fields, few studies have assessed the differential role of social media during the pandemic, by gender, from the public health perspective. Due to an overwhelming focus on women’s experiences as essential workers and with reproductive care during the COVID-19 pandemic, there is scant literature more broadly centering on women. The experiences of women were chosen as a focus in this work due to their disproportionate burden of long COVID and their higher rates of activity and gender-specific engagement patterns on social media sites [37]. This work therefore aimed to fill a gap in the extant literature by investigating the role of social media in the experiences of long-hauler women alone.

Methods

Study Design

The data used in this study were derived from an online health promotion intervention for long-hauler women with COVID-19. Participants were recruited using snowball and purposive sampling through 2 social media sites, Facebook and Slack; the participants were recruited from 16 Facebook groups and 1 Slack group, as well as 2 websites of organizations for long-hauler women. Those eligible to participate in the study met the inclusion criteria of living in the United States, being aged 18 years or older, who spoke English, and who self-identified as long-haulers due to persistent COVID-19 symptoms for 4 weeks or more after infection.

Ethical Considerations

The University of South Carolina Institutional Review Board (Pro00109358) reviewed and approved the study protocol.

XSL•FO
RenderX

Recruitment began after group and organization administrators approved posts including the study description, a flyer, and researcher contact information. The study was then advertised in each group.

Recruitment

The recruitment period spanned 2.5 months from March to June 2021. After screening for eligibility and receiving informed consent, a total of 15 semistructured, one-on-one interviews were conducted from April to June 2021 using the online videoconferencing software Zoom [38]. Each interview lasted between 30 and 50 minutes. Participant demographics were collected through the interview process. All interviews were audio-recorded and, upon completion, field notes were written. Each participant was compensated with a US $30 e-gift card for their time and effort spent participating in the study. Data saturation was agreed to have been reached, by the 2 researchers involved in interview coding, after 15 interviews.

Data Collection

Data were collected on the participants' self-reported long-hauler status, the impact of persistent COVID-19 symptoms on their lives, coping strategies, and overall experiences. In these conversations, discussions of the roles of technology and social media arose organically following the semistructured interview guide. All interviews were recorded and transcribed using the service Otter.ai [39]. All artificial intelligence–derived transcripts were reviewed and verified by members of the research team. Interviewer field notes were used as additional data.

Data Analysis

The data were analyzed following a predominately inductive approach for the thematic analysis of the interviews, as the themes identified were derived directly from the data [40-42]. The analysis process was comprised of 6 stages beginning with data familiarization and preliminary code construction, followed by the obtaining, revising, labeling, and reporting of key themes [43]. MAXQDA software was used to analyze interview transcripts [44]. In the initial phase of the thematic analysis, 2 members of the research team independently coded the transcripts, following an open coding scheme, to identify emergent themes [45-47]. The initial development of the codebook was performed after half of the interviews (ie, 7) were coded. We discussed at length the creation of the codebook to ensure accuracy of the initial codes, themes, domains, definitions, exemplar quotes, and organization. Once the codebook was finalized, the same 2 members of the research team continued to independently code the remaining transcripts. We then engaged in a collaborative review process to confirm alignment with the final codebook and to ensure consistency in the application of codes. In comparing themes, we identified similarities, differences, and interactions between themes. We used an axial coding approach to categorize the main themes and subthemes, which then guided the selection of direct quotes to demonstrate the key findings. Peer debriefing and intercoder agreement techniques were used to ensure reliability throughout the data analysis [48,49].

Results

Participant Details

The study participants, in alignment with the inclusion criteria, all identified as women. The participants were primarily aged between 36 and 65 years (n=12, 80%), served as essential workers (n=9, 60%), and lived with others (n=13, 87%) in the eastern region of South Carolina (n=10, 67%). Table 1 lists the participant details.

XSL•FO
RenderX

Table 1. Demographic characteristics of long-hauler women (N=15).

Characteristics	Participants, n (%)[a]
Age (years)	
20-35	2 (13)
36-50	6 (40)
51-65	6 (40)
>65	1 (7)
Occupation	
Health care provider	5 (33)
Educator	4 (27)
Business owner	4 (27)
Student	1 (7)
Retiree	1 (7)
Living situation	
Living with others	13 (87)
Living alone	2 (13)
Regional location	
East	10 (67)
Central	3 (20)
West	2 (13)

[a]The percentages might add up to more than 100 because of rounding.

Benefits and Challenges

Long-hauler women indicated that their most used social media features included participating in support groups, posting, commenting, connecting with others, and consuming media. They used these features of social media sites to fulfill needs such as emotional support, social engagement, spirituality, health planning, information gathering, professional support, and recreation. The different functions of social media also resulted in a variety of benefits and challenges throughout the participants' coping with long COVID. Tables 2 and 3 present the benefits and challenges related to the themes and subthemes identified and exemplar quotes.

XSL•FO
RenderX

Table 2. Emergent themes and subthemes of the beneficial roles of social media identified by long-hauler women.

Themes and subthemes	Exemplar quotes
Social connection	
Group membership	"So I said, 'Well, maybe if I am part of something…it is gonna be like a, like a, like a motivation for me to go through something for me. Um, because again, we are always thinking about others, you know, like, what, why am I complaining."
Social support	"I may not exactly know what you're going through, but I am here to help and here to listen because a lot of times you just want a hearing—somebody to hear you."
Network building	"I also was posting and doing that, which, like, kept me motivated [to continue] sharing and connecting with other people…"
Belonging	"I probably gravitated more towards that group…and I would talk about that Facebook group a lot. Like, it felt like that was like a support group, and it felt like, you know, I am not crazy, like some other people are having it, too. And I would be active in, like, commenting on, like, you know, answering people's questions or, like, sharing, like, a connection that I have with another person that wrote on there."
Religiosity and spirituality	
Prayer	"I join[ed] an online group for praying."
Fellowship	The loss of a group member highlighted a sense of duty and belongingness toward one another in the online prayer group.
Online worship	"I will go to one of my favorite pastors on YouTube and listen."
Meditation	"I am in a meditation group that I go to online, and we do meditation together. And then, there is, like, headspace and calm, those apps. So, there is a wide variety of different things. Like chakras, and then there is, you know, just all different kinds of relaxation and tension. Like, you squeeze your arms and look at your feet."
Information gathering	
Long hauler–shared information	"It is kind of, like, you form your own little support groups of people that had COVID. And, you know, their symptoms vary, and you are like, 'Oh, what did you do for this?' Or like the hair loss. That is another thing—hair loss. My hair is still not well, or whatever. And then, you know, people debating, like, 'Are you taking the vaccine? Are you not getting the vaccine?' So having those little groups to talk—it is good."
Symptom management	"I downloaded an app on my phone, and I am monitoring, like, I am documenting all of my activities for the day every day so that I can document, like, different symptoms that I am having and, like, what is, like, a trigger."
Physician-shared information	In reference to streaming YouTube videos: "…the different doctors and, like, what their findings are, what their recommendations are."
Recreation	
Entertainment	"I will allow myself; it does not happen every day, but, like, just to play some mind games, you know, a game of solitaire or a game…on my phone just to give myself a break."
Relaxation	"To help go to sleep at night. They try to, kind of, get me to relax, or whatever. And so, I think the biggest thing for me is disconnecting from all the things that I have going on, and I just…I struggle with that."

XSL•FO
RenderX

Table 3. Emergent themes and subthemes of the challenging roles of social media identified by long-hauler women.

Theme and subthemes	Exemplar quotes
Social connection	
Anxiety	"They can really increase anxiety."
Resentment	"…it was hard. There would be resentment, and there is resentment now with the group, too, as terrible as it sounds, even jealousy, because I will see people that will write on Facebook, like, 'Oh, like, I had COVID in December 2020' or 'I had COVID in January 2021,' and a part of me just, like, would hate it because it is…like you knew, like people were advising you not to travel. And that would be what it was, especially around the holiday season, hearing people talk about, you know, having so-and-so over from, like, California, and then they got sick afterwards. And, like, it just makes you go crazy because there is so much more now. But I am trying not to, like, think like that because I mean, I do not know everyone's experiences, and maybe, they really did avoid it or did their best to not get it, and they got it because we are in a pandemic."
Information gathering	
Misinformation and health literacy	"Sometimes, I can understand a lot of the stuff, but there is some things that I am not as familiar with…"
Oversaturation and pandemic fatigue	"…after a while, like, even that [social media use] got to be so overwhelming because, again, like, everyone is, like, posting the same thing."

Social Connection

A majority of the women interviewed highlighted the role of social media in reducing social isolation by providing social connection. Social connections were found to be fostered through multiple functions of social media, such as personal networks, following networks, and group membership, as well as more broadly through engagement with other users, known or unknown. Long-hauler women emphasized the importance of social media as a tool to maintain connection with their social networks when unable to be present in person. Emergent subthemes related to the main theme of social connection included the benefits of belonging fostered through social support, group membership, and network building, as well as the challenges of anxiety and resentment.

Group Membership and Network Building

One participant noted the function of social media in mediating "the loss of family time and not being able to be together and doing the things we have always done as a family." Another participant described the stress associated with physical, in-person gatherings in the time of COVID-19:

> *I tried to host a barbecue out in our little place at the lake, and it caused me so much anxiety. I could not even eat my birthday barbecue, could not really interact with people.*

At a time when minimizing physical contact with others was recommended, the online environment was found to aid in maintaining social health. Networking, a distinct feature of social media platforms, connecting individuals with others they may or may not be geographically close to, emerged as instrumental to long-hauler women's social connection and, further, social support. Participants shared motivations for seeking membership and experiences as members of online support groups for COVID-19 long-haulers. One participant noted:

> *I was looking for, you know, for common ground, for folks that were experiencing some of those same things that I was, and I was also looking to support them with what I knew about my mind, body, [and] skills…*

Facebook emerged as a popular social media platform among long-hauler women due to its functionality to host support groups. Many long-hauler women reported using Facebook groups to build their social networks, while also providing social and emotional support to other long-haulers. Upon reflecting on her participation in online social support groups, a participant shared:

> *I probably gravitated more toward that group…and I would talk about that Facebook group a lot. Like, it felt like that was like a support group, and it felt like, you know, I am not crazy, like some other people are having it, too. And I would be active in, like, commenting on, like, you know, answering people's questions or, like, sharing, like, a connection that I have with another person that wrote on there.*

Long-hauler women demonstrated the role of online groups in expanding their social networks to include other long-haulers outside their direct networks. As a result of their group membership, the majority of the participants indicated providing and receiving emotional and instrumental social support through connections fostered by membership in online support groups.

Social Support

Further, participants explained the role of online groups in facilitating social support from connections because "[they] are experiencing similar things that I am experiencing, so I know that it is not just me." One participant described her role in providing emotional social support through online social connections:

> *I may not exactly know what you are going through, but I am here to help and here to listen because a lot of times you just want a hearing—somebody to hear you.*

Participants demonstrated the crucial role of validation and affirmation as emotional social support when received from other group members regarding their emotions, symptoms, and

XSL•FO
RenderX

overall experiences. Participants indicated receiving validation and affirmation when posting, commenting, and being active within their support groups. One participant discussed the benefits of continued engagement in these groups:

> *I also was posting and doing that, which, like, kept me motivated [to continue] sharing and connecting with other people…*

As a result of providing and receiving social support within their online networks of long-hauler women, the majority of the participants report a positive effect on their sense of belonging.

Belonging

We found that because social media is able to connect users, participation on the platforms and in groups aids in maintaining social health through online belongingness, while also adhering to public health recommendations (eg, social distancing, isolation, quarantine). In bolstering social connections and aiding in emotional regulation, another long-hauler explained that social media motivates her to remain strong and encourages resilience. She explained:

> *So I said, well, maybe if I am part of something…it is gonna be like a, like a, like a motivation for me to go through something for me. Um, because again, we are always thinking about others, you know, like, what, why am I complaining.*

Another participant noted the benefit of belonging to a support group:

> *I joined the COVID long-haulers' Facebook group because another new thing with my shortness of breath is I noticed if I eat a lot at one time, I am way shorter of breath, and I do not know why. So, I was, like, "Oh, I am gonna see if anybody else has had these symptoms. So, I actually, like, made a post about it. And I liked that group because it makes you realize, like, you are not alone. There is all these other people that also do not have answers and also have similar symptoms as you.*

Participants described how their group membership and sense of belonging decreased their feelings of loneliness and isolation, particularly when sharing experiences and symptoms with other long-haulers.

Anxiety and Resentment

Despite the potential benefits of participating in online groups, there remain potential consequences of participation as well. Although the findings indicated social media aids in mitigating feelings of loneliness and creating a sense of belonging, they also indicated increasing anxiety and resentment among long-hauler women. One participant noted that "they can really increase anxiety." As related to seeing the posts of others within their social networks and in groups, a participant said:

> *…it was hard. There would be resentment, and there is resentment now with the group, too, as terrible as it sounds, even jealousy, because I will see people that will write on Facebook like "Oh, like, I had COVID in December 2020" or "I had COVID in January 2021," and a part of me just, like, would hate it because it is…like you knew, like people were advising you not to travel. And that would be what it was, especially around the holiday season, hearing people talk about, you know, having so-and-so over from, Like, California, and then they got sick afterwards. And, like, it just makes you go crazy because there is so much more now. But I am trying not to, like, think like that because I mean, I do not know everyone's experiences, and maybe, they really did avoid it or did their best to not get it, and they got it because we are in a pandemic. But it is stuff like that. Like, I feel like I am more, like, insecure with my experience. I get jealous of other people's experiences. There's just, like, a lot of negative-ness with it…*

In sharing this anecdote, the participant voiced her frustration toward and resentment of those who, after participating in high-risk activities, shared their COVID-19 experiences online. Engagement with such individuals and their posts then led to this participant's insecurity in their own experiences.

Religiosity and Spirituality

In addition to the impacts of social media on social health, participants highlighted its role in also maintaining their spiritual health. In addition to joining online groups topically centered around COVID-19, a participant indicated, "I join[ed] an online group for praying." She detailed the group, demonstrating its resemblance to that of other support groups, albeit not solely related to COVID-19, with the added element of religion. Overall, the participant's sentiments indicated that the group positively impacted her overall well-being. When describing the loss of a member of the group, she highlighted the role of fellowship and connection in the group as they lifted one another up in prayer and, in doing so, created belongingness, community, and strength.

Another participant demonstrated the role of social media as related to religiosity and spirituality by noting her use of video-streaming platforms to seek spiritual support. She described her engagement as, "I'll go to one of my favorite pastors on YouTube and listen." During a time when physically gathering with others, as in the case of congregating for religious observances, was considered high risk, social media provided an avenue through which long-haulers could maintain their spiritual practices. Relatedly, participants indicated using social media to engage in guided meditations. One shared her daughter's role in encouraging her participation:

> *She gave me some resources online, in an app, and then my daughter uses a different…she uses Spotify. So, she gave me that information, and so I kind of just went off of those suggestions, and now, I have my favorite guided meditations that I use on Spotify, and they are effective.*

Other participants said that they similarly engage in guided meditations but also participate in groups specific for meditation and relaxation. One participant described the meditation group and smartphone apps used:

XSL•FO
RenderX

> *I am in a meditation group that I go to online, and we do meditation together. And then, there is, like, headspace and calm—those apps. So, there is a wide variety of different things. Like chakras, and then there is, you know, just all different kinds of relaxation and tension. Like you squeeze your arms and look at your feet.*

Information Gathering

Apart from social networking, one of the most prominent functions of social media is the sharing of news and information. Within the context of the COVID-19 pandemic, social media served as a conduit for sharing COVID-19 news, government policies and announcements, updated prevention guidelines, and general information. The findings demonstrated that long-hauler women used social media to seek information related to COVID-19 vaccines, symptoms, and symptom management strategies, as well as to follow news related to emerging treatments.

Long Hauler–Shared Information

Regarding COVID-19 information, topics of interest were primarily related to symptoms and health management. Long-hauler women indicated turning to online support groups to gather information from those with similar experiences. One participant illustrated the symptom and health discussions within these groups:

> *It is kind of like you form your own little support groups of people that had COVID. And, you know, their symptoms vary, and you are like, "Oh, what did you do for this?" Or like the hair loss. That is another thing—hair loss. My hair is still not well, or whatever. And then, you know, people debating, like, "Are you taking the vaccine? Are you not getting the vaccine?" So having those little groups to talk—it is good.*

Alongside using support groups for discussion, long-haulers indicated also using smartphone apps to track symptoms and create health care plans. One long-hauler discussed her experience:

> *I downloaded an app on my phone, and I am monitoring, like, I am documenting all of my activities for the day every day so that I can document, like, different symptoms that I am having and, like, what is, like, a trigger.*

Due to the persistent nature of COVID-19 symptoms experienced by long-haulers, monitoring symptoms is in the interest of patients to aid in symptom management and for use with health care providers in creating treatment plans.

Physician-Shared Information

In addition to sharing information across networks of long-haulers on social media, participants also noted gathering information through online interactions with physicians and mental health professionals. One participant indicated obtaining pandemic-related information from physicians on YouTube as she watched "…the different doctors and, like, what their findings are, what their recommendations are." Due to the increasing burden of mental health challenges coupled with physical symptoms, as expressed by the participants, social media offers a platform for mental health resource sharing, at a time when many cannot access needed services. One participant detailed these difficulties:

> *I had been looking for, like, counseling, and a lot of the counseling in our plan has, like, basically stopped taking people. Like, I think it is, like, kind of like, overwhelmed right now, and, like, I would call, like, a whole list, and I would go through the whole list, and, like, they are not taking new patients. So, I just have to be persistent about it.*

In coping with barriers (ie, wait lists, cost) to accessing mental health services, participants indicated using social media as a tool to gain information from professionals. For instance, a participant said that she "join[ed] a group…they had a list of faculty members that were starting groups…you did not have to pay for it." This participant was able to engage in mental health services through a free and accessible online group operated by mental health professionals. This function of social media is valuable in responding to increasing mental health needs by addressing barriers to accessing professional psychological support.

Misinformation and Health Literacy

The potential consequences of users obtaining information from social media, particularly that which must be scientifically based, include a lack of or difficulty in understanding, as well as the distribution of and access to unvalidated content or misinformation. Illustrating the difficulty in understanding sought-out information, a participant shared:

> *Sometimes I can understand a lot of the stuff, but there is some things that I am not as familiar with…*

Due to the evolving nature of scientific discovery over the course of the pandemic, there were difficulties in grasping timelines and emergent findings that inhibited understanding and perpetuated misunderstanding. In the case of long-haulers, their increased need for health care exposes them to complex medical jargon that may require a higher level of health literacy to mitigate misunderstanding. Overall, due to the need for regularly updated COVID-19 information, social media functions as both a benefit and a hindrance to its dissemination. Social media provides users with increased access to information, while also providing a platform through which misinformation may be widely shared.

Oversaturation and Pandemic Fatigue

Further, despite the benefits of engaging in support groups and accessing pandemic-related information online, participants indicated differing perspectives on the amount of information shared. Referencing a long-hauler Facebook support group, a participant noted:

> *And, like, the nice thing about it is they share loads of information.*

Alternatively, another participant shared that due to the sameness and sheer volume of pandemic-related content on social media:

XSL•FO
RenderX

> *…after a while, like, even that [social media use] got to be so overwhelming because, again, like everyone is, like, posting the same thing.*

Due to oversaturation and misinformation, a participant noted that she is "disappointed with social media." This disappointment has kept the participant from participating in COVID-19 and long-hauler groups.

Recreation

In addition to the networking and information-gathering functionalities of social media, participants also indicated leveraging social media for entertainment, recreation, and relaxation. In coping with their diagnosis, symptoms, anxiety, and the state of the world, long-hauler women indicated using social media and smartphone apps to play games, watch videos, and listen to music. One participant described consuming content on social media as a method to cope with the anxiety of attending post–COVID-19 appointments. Another participant shared:

> *I will allow myself; it does not happen every day, but, like, just to play some mind games, you know, a game of solitaire or a game…on my phone just to give myself a break.*

Another participant indicated using the social media site YouTube as a way "to help go to sleep at night."

> *They try to, kind of, get me to relax or whatever. And so, I think the biggest thing for me is disconnecting from all the things that I have going on, and I just…I struggle with that.*

These findings suggest that social media is a method by which participants seek entertainment, recreate, and relax. These functions serve as social media–based coping mechanisms to alleviate mental health burdens.

Discussion

Principal Findings

Long-hauler women identify engagement in online support groups to be a primary use of social media during the COVID-19 pandemic. These groups are typically disease specific and can be described as communities where individuals can congregate and engage in broader group discussions as a form of social connection [50]. Support groups function to allow members to affirm their long-hauler identity, maintain connections, combat isolation, seek support, compare experiences, share remedies, and coruminate [51-53]. Long-hauler women seek reassurance through channels of connection with others who share their disease-specific identity and to cope with a social environment characterized by mortality, unemployment, resource loss, and psychological burdens of prevention measure adherence and disease [54].

Online support groups have been previously assessed in various disease contexts in the literature. A systematic review of the role of online support groups for patients with prostate cancer found that the groups not only aided in participant decision-making through their dissemination and exchange of information but also provided participants with social support [55]. A review of support groups for patients with breast cancer demonstrated that the benefits or consequences of participation in social support groups are inconclusive [56]. A systematic review of studies assessing the impacts of social support groups on patients with chronic conditions found that they demonstrate a wide array of support group implementation and outcome measurements that complicate their use in the context of the COVID-19 pandemic [57]. Within the context of COVID-19, online support groups act as a tool to comply with social distancing guidance, while maintaining connections and combating isolation, depression, and anxiety [51,52]. Additionally, a systematic review of COVID-19–specific social support groups demonstrated that although they are effective in addressing participants' psychological and psychosocial needs, due to their responsiveness to the emerging needs and challenges faced by participants, there remains a need for further research [58].

This study presented both benefits and challenges associated with the participation of long-hauler women in COVID-19–specific social support groups. The benefits include the validation of shared experiences, decreased isolation, and motivation to pursue symptom management and recovery. The challenges for long-hauler women's participation include experiences of increased anxiety due to rumination within the groups, resentment and jealousy due to others' posts of unsafe pandemic activities or recovery, and an insecurity of experiences as a result of comparison. Additionally, there is the complication of pandemic fatigue, as instigated by the overwhelming amount of posts within social support groups. Within the extant literature, overexposure to pandemic news may act as a disaster stressor that acts as a risk factor for negative psychosocial outcomes [59]. As demonstrated through these findings, related to social support and network building, social media presents an opportunity for individuals to receive support and engagement with others, while also facing potential, associated challenges.

Additional engagement on social media revolved around spirituality, entertainment, recreation, and relaxation. Beyond disease-specific support groups, long-hauler women reported relying on groups that specifically serve to maintain spiritual health. Digital media, more broadly, allows long-hauler women to engage in spiritual practices alone or with others, as desired. These novel functions are significant as spiritual health has been identified as a key coping mechanism to facilitate resilience [60]. Further, digital media has demonstrated its usefulness in the coping of long-hauler women, as they noted its use for entertainment, recreation, and relaxation through audio and visual content. The emerging pandemic literature has sought to assess the complex benefits and consequences of media usage, motivations, stress, and psychosocial outcomes that have been found to differ by demographics [61]. Despite the complex mechanisms of coping within the literature, long-hauler women in this study identified digital and social media used for entertainment to be a positive coping strategy.

In addition to the features of social media facilitating coping, long-hauler women also relied on networking sites to access pandemic-related information. One unique feature of social media is the unprecedented speed with which information can

XSL•FO
RenderX

be shared, particularly evolving pandemic information, but it also presents the risk of misinformation and associated difficulties in mitigating its negative impacts [62,63]. A key consequence, due to the nature of social media, is the tendency for information that sparks outrage, typically containing misinformation, to move the most quickly through social channels, likely stifling needed, correct information [63]. Therefore, as outrage impacts the visibility of trending topics, depending on the content, it can alter individuals' risk perceptions [63]. This study corroborated these trends due to participants' disappointment with the distribution of information and with social media overall. Additionally, when evaluating what constitutes misinformation, it is necessary to consider nuances in the perspectives of various key players (eg, patients, providers, scientists), as well as their potential contributions to the knowledge base (eg, symptomology, diagnostic criteria).

As social media content allows misinformation to trend due to its outrage-evoking characteristics, the COVID-19 pandemic is seen as syndemic with an infodemic. Within the infodemic, long-hauler women expressed experiencing difficulty in understanding health information presented online as related to vaccines, symptom management, and news. In addition to the threats posed by misinformation to public health prevention efforts, social media presents users with a plethora of information that operates to mitigate the associated negative effects [33]. Social and cultural factors influencing the perceptions of and responses to health information and risk communication are related to personal control, uncertainty, trust in institutions, and trust in media, as well as an overall sense of immediacy [63].

Associated with COVID-19 information, long-hauler women turn to social media and online social support groups to discuss symptom and health management, as well as the intention to vaccinate. Digital media, beyond social media, has benefited women with chronic COVID-19, allowing them to document and track their symptoms. As chronic COVID-19 is characterized by persistent symptoms, symptom management support and remedy information sharing were found to be salient uses of social media among women with long COVID. This finding aligns with evidence within the extant literature where social media has been used throughout the pandemic, by broader populations, to share medication strategies, anonymously seek information, crowdsource information, and engage in advocacy [64-67]. Overall, social media is used by long-hauler women to cope, exchange social support, maintain spirituality, and seek entertainment, while also disseminating information relevant to the long-hauler experience.

Strengths and Limitations

Our findings are in alignment with "uses and gratifications theory," which posits motivations for social media use as revolving around meeting certain needs, including social connection, knowledge, and relaxation, among others [68,69]. This study contributes to the sparce, evolving literature, with findings focusing on the social media usage of long-hauler women specifically.

Our study is also subject to several limitations. First, there were constraints on the analysis due to a small sample size with limited demographic variability. Second, the structure of the questions asked restricted our ability to identify patterns of usage by platform, relying, rather, on broader trends. Additionally, as support groups were used for recruitment, the findings may not be representative of the experiences of women not engaged on social media or in online support groups.

Public Health Implications

Although this study is additive to the evolving literature, strengthening the present evidence base beyond quantitative, descriptive analyses that do not account for gendered experiences, it demonstrates a need for further research. Future deductive work should consider, concurrently, comparing the uses of social media across the spectrum of gender and age based on known differences in usage. Due to the reliance on social media platforms to gather knowledge, further work is necessitated to evaluate the content and quality of information shared within online discussions and support groups. Future research should use an intersectional framework to assess the role of social media across a variety of additional identities women hold (eg, race/ ethnicity, preexisting conditions, socioeconomic status).

Conclusion

The findings of this study support the development of gender-tailored health promotion interventions that leverage the benefits of social media, while mitigating the consequences, for women with chronic COVID. As social media serves as a pandemic mitigation tool, there is a need to better understand patterns and experiences of usage [70-72]. Informed by our findings, long-hauler women should be met where they are, through the platforms and functions that they currently use, in order for public health interventions to aid them in managing long COVID and its associated effects.

Acknowledgments

This work was supported by the National Institute of Allergy and Infectious Diseases of the National Institutes of Health (grant award number R01AI127203-5S1).

Conflicts of Interest

None declared.

References

XSL•FO
RenderX

1. Yakut ZAC, Sakarcan S, Sener G. What we know about COVID-19 and its treatment. J Res Pharm 2020 Sep 11;24(5):602-616. [doi: 10.35333/jrp.2020.215]
2. Baloch S, Baloch MA, Zheng T, Pei X. The coronavirus disease 2019 (COVID-19) pandemic. Tohoku J Exp Med 2020 May;250(4):271-278 [FREE Full text] [doi: 10.1620/tjem.250.271] [Medline: 32321874]
3. COVID-19 and your health. Centers for Disease Control and Prevention. URL: https://tinyurl.com/mpj37sp8 [accessed 2024-05-01]
4. Advice for the public: coronavirus disease (COVID-19). World Health Organization. URL: https://tinyurl.com/ys6r622x [accessed 2024-05-01]
5. Coronavirus disease 2019 (COVID-19). Centers for Disease Control and Prevention. URL: https://tinyurl.com/ysa769wc [accessed 2024-05-01]
6. COVID data tracker. Centers for Disease Control and Prevention. URL: https://covid.cdc.gov/covid-data-tracker [accessed 2024-05-01]
7. United States of America: WHO coronavirus disease (COVID-19) dashboard with vaccination data. World Health Organization. URL: https://covid19.who.int [accessed 2024-07-01]
8. Bull-Otterson L, Baca S, Saydah S, Boehmer TK, Adjei S, Gray S, et al. Post–COVID conditions among adult COVID-19 survivors aged 18–64 and ≥65 years — United States, March 2020–November 2021. MMWR Morb Mortal Wkly Rep 2022 May 27;71(21):713-717. [doi: 10.15585/mmwr.mm7121e1]
9. Post-COVID conditions. Centers for Disease Control and Prevention. URL: https://tinyurl.com/yc5x7e78 [accessed 2024-07-01]
10. Wong J, Kudla A, Pham T, Ezeife N, Crown D, Capraro P, et al. Lessons learned by rehabilitation counselors and physicians in services to COVID-19 long-haulers: a qualitative study. Rehabil Couns Bull 2021 Dec 14;66(1):25-35. [doi: 10.1177/00343552211060014]
11. Del Rio C, Collins LF, Malani P. Long-term health consequences of COVID-19. JAMA 2020 Dec 03;324(17):1723-1724 [FREE Full text] [doi: 10.1001/jama.2020.19719] [Medline: 33031513]
12. Pincherle A, Jöhr J, Pancini L, Leocani L, Dalla Vecchia L, Ryvlin P, et al. Intensive care admission and early neuro-rehabilitation. Lessons for COVID-19? Front Neurol 2020 Aug 25;11:880 [FREE Full text] [doi: 10.3389/fneur.2020.00880] [Medline: 32982916]
13. Tenforde MW, Kim SS, Lindsell CJ, Billig Rose E, Shapiro NI, Files DC, CDC COVID-19 Response Team, et al. Symptom duration and risk factors for delayed return to usual health among outpatients with COVID-19 in a multistate health care systems network — United States, March-June 2020. MMWR Morb Mortal Wkly Rep 2020 Jul 31;69(30):993-998 [FREE Full text] [doi: 10.15585/mmwr.mm6930e1] [Medline: 32730238]
14. Nearly One in Five American Adults Who Have Had COVID-19 Still Have "Long COVID". Centers for Disease Control and Prevention. URL: https://www.cdc.gov/nchs/pressroom/nchs_press_releases/2022/20220622.htm [accessed 2024-10-31]
15. Bai F, Tomasoni D, Falcinella C, Barbanotti D, Castoldi R, Mulè G, et al. Female gender is associated with long COVID syndrome: a prospective cohort study. Clin Microbiol Infect 2022 May;28(4):611.e9-611.e16 [FREE Full text] [doi: 10.1016/j.cmi.2021.11.002] [Medline: 34763058]
16. Huang Y, Pinto MD, Borelli JL, Asgari Mehrabadi M, Abrahim HL, Dutt N, et al. COVID symptoms, symptom clusters, and predictors for becoming a long-hauler looking for clarity in the haze of the pandemic. Clin Nurs Res 2022 Dec 24;31(8):1390-1398 [FREE Full text] [doi: 10.1177/10547738221125632] [Medline: 36154716]
17. The sex, gender and COVID-19 project. Global Health 50/50. URL: https://tinyurl.com/2h3asj2k/ [accessed 2024-05-01]
18. Carli LL. Women, gender equality and COVID-19. Gender Manag Int J 2020 Sep 30;35(7/8):647-655. [doi: 10.1108/gm-07-2020-0236]
19. Yaksi N, Teker AG, Imre A. Long hospitalized COVID-19 patients: a retrospective cohort study. Iran J Public Health 2022 Jan 10;51(1):88-95 [FREE Full text] [doi: 10.18502/ijph.v51i1.8297] [Medline: 35223629]
20. Aiyegbusi OL, Hughes SE, Turner G, Rivera SC, McMullan C, Chandan JS, TLC Study Group. Symptoms, complications and management of long COVID: a review. J R Soc Med 2021 Oct 15;114(9):428-442 [FREE Full text] [doi: 10.1177/01410768211032850] [Medline: 34265229]
21. Torjesen I. COVID-19: middle aged women face greater risk of debilitating long term symptoms. BMJ 2021 Mar 25;372:n829. [doi: 10.1136/bmj.n829] [Medline: 33766927]
22. Mobile fact sheet. Fact sheets: tech adoption trends. Pew Research Center. URL: https://tinyurl.com/ye5yzd4d [accessed 2024-05-01]
23. Computer and internet use in the United States: 2018. U.S. Census Bureau. URL: https://tinyurl.com/5746fjpd [accessed 2024-05-01]
24. Chidiac M, Ross C, Marston HR, Freeman S. Age and gender perspectives on social media and technology practices during the COVID-19 pandemic. Int J Environ Res Public Health 2022 Oct 27;19(21):13969 [FREE Full text] [doi: 10.3390/ijerph192113969] [Medline: 36360853]
25. Parker K, Horowitz JM, Minkin R. How the coronavirus outbreak has – and hasn't – changed the way Americans work. Pew Research Center. 2020 Dec 9. URL: https://tinyurl.com/2setyh4b [accessed 2024-05-01]

XSL•FO
RenderX

26. Callard F, Perego E. How and why patients made long Covid. Soc Sci Med 2021 Jan;268:113426 [FREE Full text] [doi: 10.1016/j.socscimed.2020.113426] [Medline: 33199035]
27. Miyake E, Martin S. Long Covid: online patient narratives, public health communication and vaccine hesitancy. Digit Health 2021 Nov 29;7:20552076211059649 [FREE Full text] [doi: 10.1177/20552076211059649] [Medline: 34868622]
28. David ME, Roberts JA. Smartphone use during the COVID-19 pandemic: social versus physical distancing. Int J Environ Res Public Health 2021 Jan 25;18(3):1034 [FREE Full text] [doi: 10.3390/ijerph18031034] [Medline: 33503907]
29. Moore K, March E. Socially connected during COVID-19: online social connections mediate the relationship between loneliness and positive coping strategies. J-STAR 2022 Jan 10;1(1):1-20. [doi: 10.55319/js.v1i1.9]
30. Hajek A, König HH. The association between use of online social networks sites and perceived social isolation among individuals in the second half of life: results based on a nationally representative sample in Germany. BMC Public Health 2019 Jan 09;19(1):40 [FREE Full text] [doi: 10.1186/s12889-018-6369-6] [Medline: 30626369]
31. Changes in internet use when coping with stress: older adults during the COVID-19 pandemic. Elsevier Enhanced Reader. URL: https://pubmed.ncbi.nlm.nih.gov/32771312/ [accessed 2024-05-01]
32. Garfin DR. Technology as a coping tool during the coronavirus disease 2019 (COVID-19) pandemic: implications and recommendations. Stress Health 2020 Oct 13;36(4):555-559 [FREE Full text] [doi: 10.1002/smi.2975] [Medline: 32762116]
33. Cato S, Iida T, Ishida K, Ito A, Katsumata H, McElwain KM, et al. Social media infodemics and social distancing under the COVID-19 pandemic: public good provisions under uncertainty. Glob Health Action 2021 Jan 01;14(1):1995958 [FREE Full text] [doi: 10.1080/16549716.2021.1995958] [Medline: 34806565]
34. Clark JL, Algoe SB, Green MC. Social network sites and well-being: the role of social connection. Curr Dir Psychol Sci 2017 Dec 19;27(1):32-37. [doi: 10.1177/0963721417730833]
35. Lin X, Wang X. Examining gender differences in people's information-sharing decisions on social networking sites. Int J Inf Manag 2020 Feb;50:45-56. [doi: 10.1016/j.ijinfomgt.2019.05.004]
36. Krasnova H, Veltri NF, Eling N, Buxmann P. Why men and women continue to use social networking sites: the role of gender differences. J Strateg Inf Syst 2017 Dec;26(4):261-284. [doi: 10.1016/j.jsis.2017.01.004]
37. SPercentage of adults in the United States who use social networks as of February 2021, by gender. Statista. 2019. URL: https://www.statista.com/statistics/471345/us-adults-who-use-social-networks-gender/ [accessed 2024-05-01]
38. Video conferencing, web conferencing, webinars, screen sharing. Zoom. URL: https://us06web.zoom.us/ [accessed 2024-07-01]
39. Otter.ai. URL: https://otter.ai/ [accessed 2024-01-31]
40. Boyatzis RE. Transforming qualitative information: thematic analysis and code development. Thousand Oaks, CA: SAGE Publications; 1998:xvi, 184.
41. Fereday J, Muir-Cochrane E. Demonstrating rigor using thematic analysis: a hybrid approach of inductive and deductive coding and theme development. Int J Qual Methods 2016 Nov 29;5(1):80-92. [doi: 10.1177/160940690600500107]
42. Guest G, Bunce A, Johnson L. How many interviews are enough? Field Methods 2016 Jul 21;18(1):59-82. [doi: 10.1177/1525822X05279903]
43. Braun V, Clarke V. Using thematic analysis in psychology. Qual Res Psychol 2006 Jan;3(2):77-101. [doi: 10.1191/1478088706qp063oa]
44. All-in-one qualitative data analysis software. MAXQDA. URL: https://www.maxqda.com/qualitative-data-analysis-software [accessed 2024-07-01]
45. Burnard P. A method of analysing interview transcripts in qualitative research. Nurse Educ Today 1991 Dec;11(6):461-466. [doi: 10.1016/0260-6917(91)90009-y] [Medline: 1775125]
46. Polit DF, Beck CT. Nursing research: principles and methods. Philadelphia, PA: Lippincott Williams & Wilkins; 2004.
47. Corbin J, Strauss A. Basics of qualitative research: techniques and procedures for developing grounded theory. Fourth edition. Thousand Oaks, CA: SAGE Publications; 2014.
48. Krippendorff K. Content analysis: an introduction to its methodology. Fourth edition. Thousand Oaks, CA: SAGE Publications; 2019.
49. Padgett DK. Qualitative methods in social work research. Third edition. Thousand Oaks, CA: SAGE Publications; 2008.
50. Leist AK. Social media use of older adults: a mini-review. Gerontology 2013 Apr 16;59(4):378-384 [FREE Full text] [doi: 10.1159/000346818] [Medline: 23594915]
51. Pelletier J, Houle J, Goulet M, Juster R, Giguère CÉ, Bordet J, et al. Online and recovery-oriented support groups facilitated by peer support workers in times of COVID-19: protocol for a feasibility pre-post study. JMIR Res Protoc 2020 Dec 18;9(12):e22500 [FREE Full text] [doi: 10.2196/22500] [Medline: 33259326]
52. Brooks SK, Webster RK, Smith LE, Woodland L, Wessely S, Greenberg N, et al. The psychological impact of quarantine and how to reduce it: rapid review of the evidence. Lancet 2020 Mar 14;395(10227):912-920 [FREE Full text] [doi: 10.1016/S0140-6736(20)30460-8] [Medline: 32112714]
53. Stone LB, Veksler AE. Stop talking about it already! Co-ruminating and social media focused on COVID-19 was associated with heightened state anxiety, depressive symptoms, and perceived changes in health anxiety during spring 2020. BMC Psychol 2022 Mar 07;10(1):22 [FREE Full text] [doi: 10.1186/s40359-022-00734-7] [Medline: 35130965]

XSL•FO
RenderX

54. Lee SA, Jobe MC, Mathis AA, Gibbons JA. Incremental validity of coronaphobia: coronavirus anxiety explains depression, generalized anxiety, and death anxiety. J Anxiety Disord 2020 Aug;74:102268 [FREE Full text] [doi: 10.1016/j.janxdis.2020.102268] [Medline: 32650221]
55. Ihrig A, Karschuck P, Haun MW, Thomas C, Huber J. Online peer-to-peer support for persons affected by prostate cancer: a systematic review. Patient Educ Couns 2020 Oct;103(10):2107-2115. [doi: 10.1016/j.pec.2020.05.009] [Medline: 32475711]
56. McCaughan E, Parahoo K, Hueter I, Northouse L, Bradbury I. Online support groups for women with breast cancer. Cochrane Database Syst Rev 2017 Mar 10;3(3):CD011652 [FREE Full text] [doi: 10.1002/14651858.CD011652.pub2] [Medline: 28278559]
57. Thompson DM, Booth L, Moore D, Mathers J. Peer support for people with chronic conditions: a systematic review of reviews. BMC Health Serv Res 2022 Mar 31;22(1):427 [FREE Full text] [doi: 10.1186/s12913-022-07816-7] [Medline: 35361215]
58. Margherita G, Muzii B, Caffieri A, Di Francia A, Somma B. 'Isolated together': online group treatments during the COVID-19 pandemic. A systematic review. Res Psychother 2022 Dec 04;25(3):639 [FREE Full text] [doi: 10.4081/ripppo.2022.639] [Medline: 36373389]
59. Zhao N, Zhou G. Social media use and mental health during the COVID-19 pandemic: moderator role of disaster stressor and mediator role of negative affect. Appl Psychol Health Well Being 2020 Dec 17;12(4):1019-1038 [FREE Full text] [doi: 10.1111/aphw.12226] [Medline: 32945123]
60. Roberto A, Sellon A, Cherry ST, Hunter-Jones J, Winslow H. Impact of spirituality on resilience and coping during the COVID-19 crisis: a mixed-method approach investigating the impact on women. Health Care Women Int 2020 Oct 21;41(11-12):1313-1334. [doi: 10.1080/07399332.2020.1832097] [Medline: 33085579]
61. Pahayahay A, Khalili-Mahani N. What media helps, what media hurts: a mixed methods survey study of coping with COVID-19 using the media repertoire framework and the appraisal theory of stress. J Med Internet Res 2020 Aug 06;22(8):e20186 [FREE Full text] [doi: 10.2196/20186] [Medline: 32701459]
62. Venegas-Vera AV, Colbert GB, Lerma EV. Positive and negative impact of social media in the COVID-19 era. Rev Cardiovasc Med 2020 Dec 30;21(4):561-564 [FREE Full text] [doi: 10.31083/j.rcm.2020.04.195] [Medline: 33388000]
63. Malecki K, Keating JA, Safdar N. Crisis communication and public perception of COVID-19 risk in the era of social media. Clin Infect Dis 2021 Feb 16;72(4):697-702 [FREE Full text] [doi: 10.1093/cid/ciaa758] [Medline: 32544242]
64. Thompson CM, Rhidenour KB, Blackburn KG, Barrett AK, Babu S. Using crowdsourced medicine to manage uncertainty on Reddit: the case of COVID-19 long-haulers. Patient Educ Couns 2022 Feb;105(2):322-330 [FREE Full text] [doi: 10.1016/j.pec.2021.07.011] [Medline: 34281723]
65. Koss J, Bohnet-Joschko S. Social media mining of long-COVID self-medication reported by Reddit users: feasibility study to support drug repurposing. JMIR Form Res 2022 Oct 03;6(10):e39582 [FREE Full text] [doi: 10.2196/39582] [Medline: 36007131]
66. De Choudhury M, Morris MR, White RW. Seeking and sharing health information online: comparing search engines and social media. 2014 Presented at: CHI '14: SIGCHI Conference on Human Factors in Computing Systems; April 26-May 1, 2014; Toronto, Canada p. 1365-1376. [doi: 10.1145/2556288.2557214]
67. Pavalanathan U, De Choudhury M. Identity management and mental health discourse in social media. Proc Int World Wide Web Conf 2015 May;2015(Companion):315-321 [FREE Full text] [doi: 10.1145/2740908.2743049] [Medline: 27376158]
68. Katz E, Haas H, Gurevitch M. On the use of the mass media for important things. Am Sociol Rev 1973 Apr;38(2):164. [doi: 10.2307/2094393]
69. Korhan O, Ersoy M. Usability and functionality factors of the social network site application users from the perspective of uses and gratification theory. Qual Quant 2015 Jun 14;50(4):1799-1816. [doi: 10.1007/s11135-015-0236-7]
70. Hajek A, König HH. Social isolation and loneliness of older adults in times of the COVID-19 pandemic: can use of online social media sites and video chats assist in mitigating social isolation and loneliness? Gerontology 2021;67(1):121-124 [FREE Full text] [doi: 10.1159/000512793] [Medline: 33264778]
71. YU RP, MCCAMMON RJ, ELLISON NB, LANGA KM. The relationships that matter: social network site use and social wellbeing among older adults in the United States of America. Ageing Soc 2015 Jul 02;36(9):1826-1852. [doi: 10.1017/s0144686x15000677]
72. Latikka R, Koivula A, Oksa R, Savela N, Oksanen A. Loneliness and psychological distress before and during the COVID-19 pandemic: relationships with social media identity bubbles. Soc Sci Med 2022 Jan;293:114674 [FREE Full text] [doi: 10.1016/j.socscimed.2021.114674] [Medline: 34959045]

XSL•FO
RenderX

Edited by A Kushniruk; submitted 30.06.23; peer-reviewed by L Laestadius, J Himstedt; comments to author 20.11.23; revised version received 15.12.23; accepted 18.02.24; published 23.04.24.

Please cite as:
Garrett C, Aghaei A, Aggarwal A, Qiao S
The Role of Social Media in the Experiences of COVID-19 Among Long-Hauler Women: Qualitative Study
JMIR Hum Factors 2024;11:e50443
URL: https://humanfactors.jmir.org/2024/1/e50443
doi:10.2196/50443
PMID:38652515

©Camryn Garrett, Atefeh Aghaei, Abhishek Aggarwal, Shan Qiao. Originally published in JMIR Human Factors (https://humanfactors.jmir.org), 23.04.2024. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on https://humanfactors.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Usability and Acceptability of a Conversational Agent Health Education App (Nthabi) for Young Women in Lesotho: Quantitative Study

Elizabeth Nkabane-Nkholongo[1*], RN, MPH; Mathildah Mpata-Mokgatle[1*], PHD; Brian W Jack[2*], MD; Clevanne Julce[3*], MPH; Timothy Bickmore[4*], PHD

[1]School of Public Health, Sefako Makgatho Health Sciences University, Pretoria, South Africa
[2]Chobanian & Avedisian School of Medicine, Boston University, Boston, MA, United States
[3]Umass Chan Medical School, University of Massachusetts, Worcester, MA, United States
[4]Khoury College of Computer Sciences, Northeastern University, Boston, MA, United States
[*]all authors contributed equally

Corresponding Author:
Elizabeth Nkabane-Nkholongo, RN, MPH
School of Public Health
Sefako Makgatho Health Sciences University
Motlalegi Street
Garankuwa
Pretoria, 0208
South Africa
Phone: 27 (012) 521 5759 ext 012
Fax: 27 (012) 560 0086
Email: elizabeth.nkholongo@gmail.com

Abstract

Background: Young women in Lesotho face myriad sexual and reproductive health problems. There is little time to provide health education to women in low-resource settings with critical shortages of human resources for health.

Objective: This study aims to determine the acceptability and usability of a conversational agent system, the Nthabi health promotion app, which was culturally adapted for use in Lesotho.

Methods: We conducted a descriptive quantitative study, using a 22-item Likert scale survey to assess the perceptions of the usability and acceptability of 172 young women aged 18-28 years in rural districts of Lesotho, who used the system on either smartphones or tablets for up to 6 weeks. Descriptive statistics were used to calculate the averages and frequencies of the variables. χ^2 tests were used to determine any associations among variables.

Results: A total of 138 participants were enrolled and completed the survey. The mean age was 22 years, most were unmarried, 56 (40.6%) participants had completed high school, 39 (28.3%) participants were unemployed, and 88 (63.8%) participants were students. Respondents believed the app was helpful, with 134 (97.1%) participants strongly agreeing or agreeing that the app was "effective in helping them make decisions" and "could quickly improve health education and counselling." In addition, 136 (98.5%) participants strongly agreed or agreed that the app was "simple to use," 130 (94.2 %) participants reported that Nthabi could "easily repeat words that were not well understood," and 128 (92.7%) participants reported that the app "could quickly load the information on the screen." Respondents were generally satisfied with the app, with 132 (95.6%) participants strongly agreeing or agreeing that the health education content delivered by the app was "well organised and delivered in a timely way," while 133 (96.4%) participants "enjoyed using the interface." They were satisfied with the cultural adaptation, with 133 (96.4%) participants strongly agreeing or agreeing that the app was "culturally appropriate and that it could be easily shared with a family or community members." They also reported that Nthabi was worthwhile, with 127 (92%) participants reporting that they strongly agreed or agreed that they were "satisfied with the application and intended to continue using it," while 135 (97.8%) participants would "encourage others to use it." Participants aged 18-24 years (vs those aged 25-28 years) agreed that the "Nthabi app was simple to use" (106/106, 100% vs 30/32, 98.8%; P=.01), and agreed that "the educational content was well organised and delivered in a timely way" (104/106, 98.1% vs 28/32, 87.5%; P=.01).

XSL•FO
RenderX

Conclusions: These results support further study of conversational agent systems as alternatives to traditional face-to-face provision of health education services in Lesotho, where there are critical shortages of human resources for health.

Trial Registration: ClinicalTrials.gov NCT04354168; https://www.clinicaltrials.gov/study/NCT04354168

(JMIR Hum Factors 2024;11:e52048) doi:10.2196/52048

KEYWORDS

preconception care; conversational agent technology; women's health education; mHealth adaptation; health information technology; health education in Africa; education; women's health; women; woman; health information; young women; survey; usability; acceptability; application; applications; app; health promotion

Introduction

Background

Digital health interventions offer considerable promise to develop new models of health care delivery and to have a large public health impact [1]. Digital channels, such as the internet, mobile phone messaging, social media, apps, voice video messaging, and telemedicine have been shown to improve the delivery of health education and care. These tools have tremendous potential to impact large-scale health promotion efforts as a cost-effective and scalable solution to address public health challenges, such as delivering sexual health education [2].

The rapid diffusion of mobile technology and advances in artificial intelligence have facilitated this trend [3]. The use of mobile devices and services has continued to increase globally, though at different rates in developed and developing countries. By the end of 2018, more than 5 billion people worldwide subscribed to mobile services, accounting for 67% of the global population, and this number is expected to exceed 70% by 2025 [4].

In Lesotho, 94% of people aged 18-29 years use smartphones, and 3G data coverage is available in almost 90% of the country [5]. This is an important group to target, as they represent the highest proportion of global consumers of mobile technology. This high penetration of mobile technologies provides an opportunity to assess the usability and acceptability of using new mobile health technologies as an alternative to the traditional face-to-face provision of health education.

Adolescents and young women continue to report low levels of sexual and reproductive health knowledge, and engage in risky sexual behaviors [6]. They also face a myriad of sexual and reproductive health problems, such as unplanned pregnancy, sexually transmitted infections, and HIV infections. Advancing sexual and reproductive health education for adolescents and young women in Africa is particularly important, as HIV accounts for 42% of new HIV infections globally [7], and 4 in 5 young people with HIV live in sub-Saharan Africa [8]. Therefore, developing new ways to provide sexual and reproductive health education in Africa is particularly important.

Lesotho is a lower middle-income country in southern Africa and has the second highest HIV prevalence in the world—at 22.7%—and one of the highest HIV incidences among adolescent girls and young women (0.33%) [9]. The maternal mortality ratio in Lesotho is the second highest in Southern African Development Community countries (544/100,000 live births) [10]. The ratio of doctors to the population is 0.9 per 10,000. For nurse-midwives, the ratio is 10.2 per 10,000, [11] which poses a challenge to the delivery of face-to-face health education.

Delivering health education via new mobile health tools has the potential to provide alternatives to traditional face-to-face provision of health education. Conversational agents are computer-based animated characters that are designed to simulate face-to-face human interactions. The human–computer interface relies only minimally on text comprehension and prioritizes conversation, thereby making it more accessible to patients with limited health literacy [12]. In health care, patient-facing conversational agents are increasingly used to deliver education, provide self-management of chronic conditions, perform routine tasks, such as appointment booking, and support health professionals' decision-making for diagnosis and triage in mental health [13,14]. These devices have the potential to automate tasks, improve access to health care services, and reduce health professionals' workload.

Prior Work

In the United States, a conversational agent named Gabby was designed to deliver preconception sexual and reproductive health information to reproductive-age African American women. Using Gabby demonstrated significant improvement in addressing reproductive health risks in randomized controlled trials [15,16].

Our research team culturally adapted Gabby to provide sexual and reproductive health education to young women in Lesotho. The newly adapted system, named the Nthabi Preconception Health Promotion App (hereafter referred to as Nthabi) is a patient-facing conversational agent that screens for sexual and reproductive health risks, and uses behavior change techniques, such as motivational interviewing and shared decision-making, to facilitate behavior change related to these risks.

The perceived appropriateness of Nthabi adaptation was studied in focus groups with young women aged 18-28 years (n=33 participants) who had used the system for 4 weeks [17]. Participants reported that adaptations were culturally appropriate, and provided relevant and culturally sensitive clinical information. They emphasized that the physical characteristics, personal and nonverbal behaviors, use of Sesotho (the local language in Lesotho) words and idioms, and clinical content were sensitively delivered and culturally appropriate. Interviews with the Ministry of Health key informants agreed that the adaptation was successful and that the system holds

XSL•FO
RenderX

great potential to improve the delivery of health education content in Lesotho.

Goal of This Study

The goal of this study is to assess the perceived usability and acceptability of the Nthabi Preconception Health Promotion App among 160 young women enrolled in a clinical trial in Lesotho who had used the system for up to 6 weeks.

Methods

Study Design

In this paper, we report the results of a survey designed to assess the perceived usability and acceptability of Nthabi among the first 160 young women who used the system.

Usability is defined as the extent to which young women can use Nthabi to achieve specific goals with effectiveness, efficiency, and satisfaction [18]. Acceptability includes the satisfaction of the young women, attitudes toward using the app, and intention or willingness to continue using the app.

Study Population and Setting

The population studied was young women aged 18 to 28 years in the Leribe and Berea districts of the rural, mountainous, lower middle-income country of Lesotho in southern Africa.

Sampling

This study was conducted to assess the usability and acceptability of using Nthabi as a health education tool in Lesotho; therefore, a convenience sample of 200 young women was chosen from the population of young women in the districts of Leribe and Berea.

Recruitment

Participants were recruited in several ways. First, the research team posted messages on social media (eg, WhatsApp and Facebook) that described the study and asked potential participants to contact the research team to discuss enrolling in the study. A nongovernmental organization called Help Lesotho, which offers mentorship programs to adolescent girls and young women in the Leribe district, saw the social media posting, reached out to the research team, and offered to disseminate the recruitment announcement to their clients.

Second, the research team directly approached young women while they were waiting for consultation at the Adolescent Health Corners (clinics) and HIV and Mother and Child Health ambulatory clinical departments at the Berea and Leribe government district hospitals. Last, students were approached at the Leribe Vocational High School and the Limkokwing University of Technology to identify individuals who might be interested in participating.

Eligibility Criteria

The inclusion criteria were the following: (1) Basotho women aged 18-28 years who were from the districts of Leribe and Berea and accessed health services in these 2 districts, (2) self-reported ability to read and understand spoken English, (3) access to an Android smartphone, and (4) ability to access internet and Wi-Fi at least once at the end of the study. Those not meeting these criteria were excluded.

Enrollment

The research team assisted the participants in downloading the app on their mobile phones. Participants who were unable to download the app on their mobile phones were loaned a Lenovo Android 11 OS platform tablet to use for 6 weeks. Participants were then assisted to create a unique username and password and were shown how to log on to either their Android mobile phone or tablet and start interacting with Nthabi. Participants were encouraged to use the app at least once daily at their convenience for 6 weeks.

Baseline Data Collection

Sociodemographic information was collected (age, marital status, education level, employment status, recruitment site, and district). A total of 160 participants were enrolled. Participant contact information (phone and WhatsApp number, email address) was collected so they could be reminded to return to the recruitment site so they could access the internet when they were finished using Nthabi, to facilitate survey completion, and return the loaned tablets.

Description of the Nthabi Intervention

Nthabi was adapted in relation to physical characteristics, language, culture, and clinical content appropriate for Lesotho, as previously described (Figure 1) [17]. A description of Nthabi is found in Multimedia Appendix 1.

Briefly, Nthabi is an English-speaking Mosotho (person from Lesotho) nurse-midwife dressed as a professional nurse. Her hairstyle (braids), complexion (medium, similar to the local population), facial expressions (calm and gentle), and mannerisms (a humble professional with a sense of humor) were relatable to young women in Lesotho.

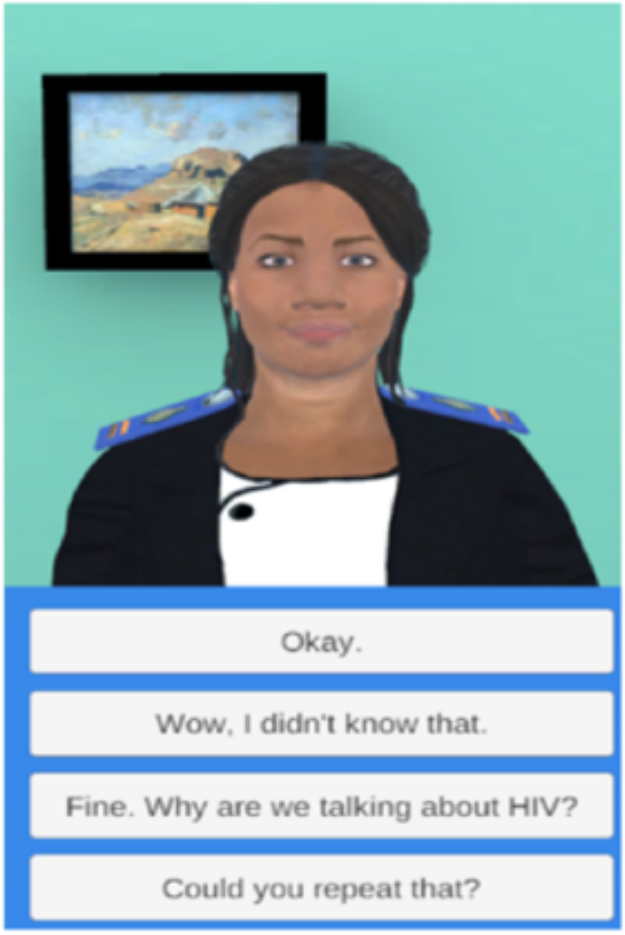
To establish the clinical topics to be included in the system, Ministry of Health key informants recommended 5 sexual reproductive health topics for young women (family planning, HIV, tuberculosis, healthy eating, and using folic acid). The research team then used the Lesotho National Clinical Guidelines on these topics to create evidence-based dialogue for use in Nthabi interactions.

During subsequent interactions with Nthabi, women selected the topic they wanted to discuss. Using conversational dialogue, Nthabi describes why the topic is important and offers suggestions about how to take action on it. The woman engages with the app by selecting a response from a multiple-choice menu that is updated at each turn of dialogue.

To increase the accessibility and use of the system, a decision was made that the app would be fully downloadable to the user's mobile phone, thereby enabling full content availability beyond the Wi-Fi environment. Use and information about the content discussed would be downloaded when the user was in a Wi-Fi environment. Nthabi was available from the Google Play store for downloading on mobile phones or tablets.

XSL•FO
RenderX

Figure 1. Nthabi health education interface.



Data Collection Tool

The survey instrument was based on the System Usability Scale and the Mobile App Rating Scale [19] using previous studies of Gabby adaptation [20,21] and modified for use in Lesotho. To ensure that the questions were clear and not ambiguous, the survey tool was reviewed by 12 health professionals, including nurses working in adolescent health, physicians, and district sexual reproductive health clinicians. The survey was then piloted with young women who met the eligibility criteria, to assess the respondents' understanding and interpretation of the questions. Only editorial changes, to enhance clarity, were required. The final survey contained 22 questions that elicited responses on a 4-point Likert scale (strongly agree, agree, disagree, and strongly disagree). Topics covered in the survey were usability (ease of use and reliability), satisfaction, willingness to continue, how easy it was to understand, content organization, and cultural relevance. The survey also enquired about the degree to which Nthabi helped women make health decisions and the degree to which they would encourage others to use the app.

Data Storage and Analysis

Survey data were captured on an Excel (Microsoft Corp) spreadsheet and stored on a password-protected computer. Data were analyzed using Stata software (StataCorp). Descriptive statistics were used to calculate the averages and frequencies of the variables. Inferential statistics, such as χ^2 tests, were used to determine any associations among variables. Statistical significance was set as a threshold of $P<.10$, as this was a feasibility study.

Participant Incentives

All participants received 50 Maloti (approximately US $5) to cover data costs. Participants using tablets were provided an additional 50 Maloti (approximately US $5) to cover their travel back to the recruitment sites to return the devices.

Ethical Considerations

Once the eligibility of participants was confirmed, the research team explained the purpose of the study, potential risks and benefits, compensation for travel costs, and the right to withdraw from the study at any time. After questions had been addressed, participants were asked to sign an informed consent form and were enrolled.

The study was conducted according to the Consolidated Standards of Reporting Trials (CONSORT) [22] and the adaptations for mobile health interventions [23]. Ethical clearance was obtained from the Boston University Research institutional review board (H-40268), Sefako Makgatho University of Health Sciences Ethics Review Committee (SMUREC/H/343/2021: PG), and the Lesotho Ministry of Health Research Ethics Committee (ID 145-2021). Permission was obtained from the study recruitment sites.

XSL•FO
RenderX

Results

Recruitment

The research team screened 436 young women for eligibility, as shown in the CONSORT diagram (Figure 2). Young women were recruited through social media (eg, WhatsApp and Facebook) or direct contact at Limkokwing University of Technology (n=150), Leribe Vocational School (n=88), Leribe Health Facilities (n=55), Berea Health Facilities (n=84), and Help Lesotho (n=59).

Of those screened, 174 young women were ineligible due to having smartphones without the Android operating system, while 64 young women had phones that were not smartphones, and 10 young women had Huawei Android smartphones that lacked access to the Google Play store.

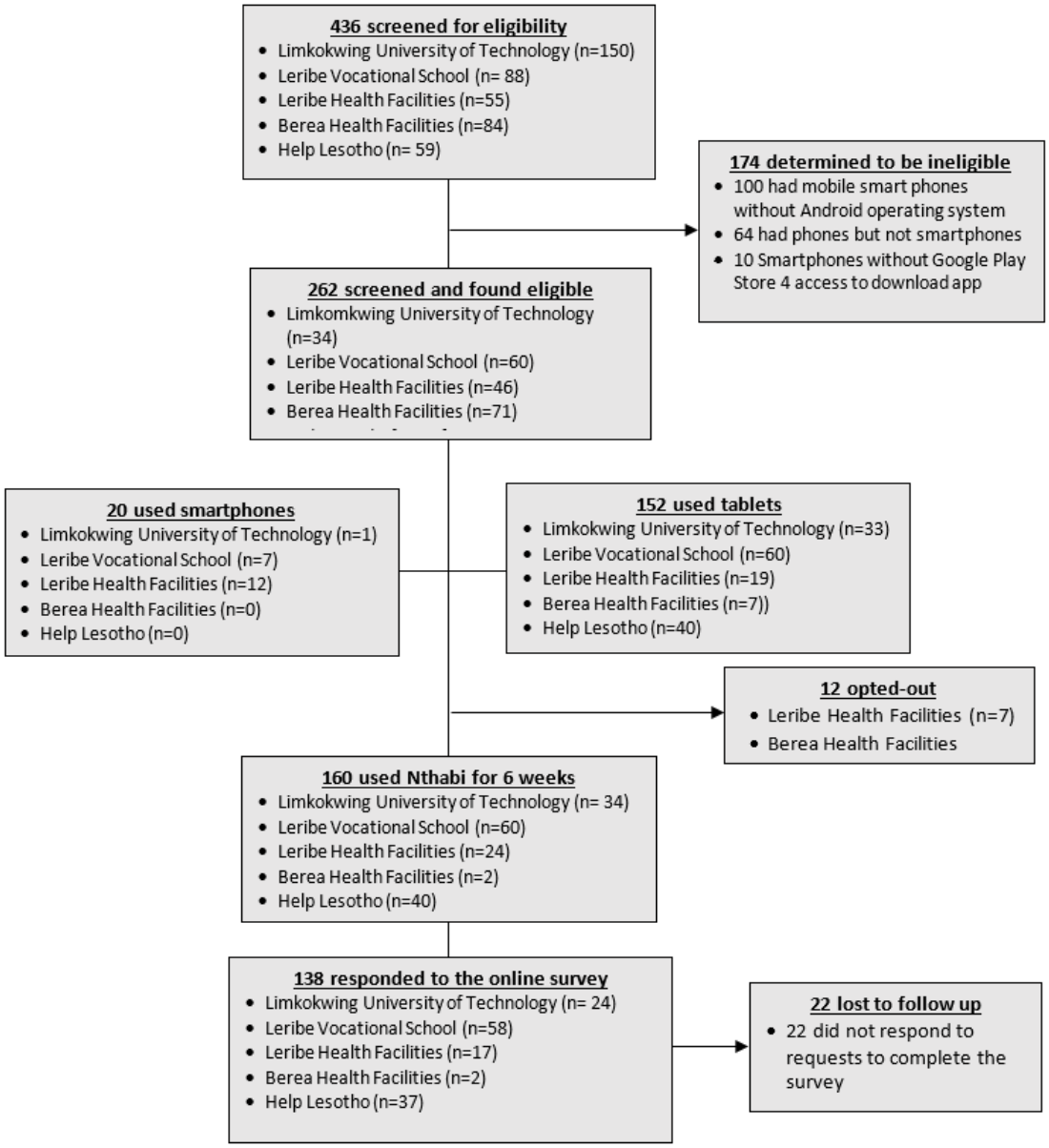
Consequently, 172 participants were eligible, provided consent, and were enrolled. Those enrolled were from Limkokwing University of Technology (34 of 34 screened), Leribe Vocational School (60 of 60 participants screened), Leribe Health Facilities (31 of 46 participants screened), Berea Health Facilities (7 of 71 participants screened), and Help Lesotho (40 of 51 participants screened).

Of those enrolled, only 20 participants had sufficient memory on their phones to download the Nthabi app, and 152 participants received a tablet device to use. Of those who were able to download the Nthabi app on their mobile phones, 1 of 34 participants was from Limkokwing University of Technology, 7 of 60 participants were from Leribe Vocational School, and 12 of 31 participants were from Leribe Health Facilities.

In the weeks after enrollment, 12 participants opted out of the study because their phones froze and jammed when they tried to load the app. Therefore, 160 young women used Nthabi for up to 6 weeks, with 8 young women using phones and 152 young women using loaned tablets. At the end of 6 weeks, 138 young women responded to the survey (80 young women who had been recruited from the technology and vocational schools, 19 young women from the district health facilities, and 37 young women from the Help Lesotho program), and 22 young women did not respond to requests to complete the survey.

XSL•FO
RenderX

Figure 2. Consolidated Standards of Reporting Trials (CONSORT) diagram.



Sociodemographic Results

Table 1 shows the characteristics of the 138 participants who were enrolled and who completed the survey after 6 weeks. The mean age was 22 years (SD 2.7 years), most were unmarried, 56 (40.6%) participants had completed high school, 39 (28.3%) participants were unemployed, and 88 (63.8%) participants were students. The recruitment sites of those participants completing surveys were 24 (17.4%) participants from Limkokwing University of Technology, 58 (42%) participants from Leribe Vocational School, 17 (12.3%) participants from Leribe health facilities, 2 (1.4%) participants from Berea Health Facilities, and 37 (26.8%) participants from Help Lesotho.

XSL•FO
RenderX

Table 1. Sociodemographic characteristics of respondents (n=138).

Characteristics	Respondents, n (%)
Age (years)	
18-20	34 (24.6)
21-23	53 (38.4)
24-26	37 (26.8)
27-28	14 (10.1)
Marital status	
Married	11 (8)
Not married	127 (92)
Level of education	
Primary	2 (1.5)
High school	56 (40.6)
College	39 (28.3)
University	41 (29.7)
Employment status	
Employed	11 (8)
Unemployed	39 (28.3)
Student	88 (63.8)
Recruitment site	
Limkokwing University of Technology	24 (17.4)
Leribe Vocational School	58 (42)
Leribe Health Facilities	17 (12.3)
Berea Health Facilities	2 (1.4)
Help Lesotho	37 (26.8)

Survey Results

Table 2 shows the survey responses of the 138 young women who completed the survey. Overall, the results show that participants perceived usability and acceptability positively.

Described below are the survey responses corresponding to the components of usability (effectiveness, efficiency, and satisfaction) and acceptability (satisfaction, attitudes toward use, and intention to continue using Nthabi).

Respondents believed the app was helpful, with 134 (97.1%) participants strongly agreeing or agreeing that the app was “effective in helping them make decisions” and “could quickly improve health education and counselling.”

Participants generally liked using the app, with 136 (98.6%) participants strongly agreeing or agreeing that the app was “simple to use,” while 132 (95.7%) participants reported that “symbols and buttons are easy to use,” 130 (94.3%) participants reported that Nthabi could “easily repeat words that were not well understood,” and 128 (92.8%) participants reported that the app “could quickly load the information on the screen.”

Respondents were generally satisfied with the app, with 132 (95.7%) participants strongly agreeing or agreeing that the health education content delivered by the app was “well organised and delivered in a timely way,” while 133 (96.4%) participants “enjoyed using the interface.”

In addition, 132 (95.7%) participants strongly agreed or agreed that they were able to complete tasks quickly using the app, while 136 (98.6%) participants reported that “I can quickly remember how to use the app after a while,” and 137 (99.3%) participants reported that “it was easy to learn how to use the app.”

The items rated less positively include the following: “it was easy to converse and type responses into the app” according to 95 (68.8%) participants, and “I could easily correct mistakes” according to 106 (76.8%) participants.

They also were satisfied with the cultural adaptation, with 133 (96.4%) participants strongly agreeing or agreeing that the app was “culturally appropriate and that it could be easily shared with a family or community members.”

Finally, they also reported that Nthabi was worthwhile, with 127 (92%) participants reporting that they strongly agreed or agreed that they were “satisfied with the application and intended to continue using it” while 135 (97.8%) participants would “encourage others to use it.”

XSL•FO
RenderX

Table 2. Survey responses on the usability and acceptability of Nthabi app (n=138).

To what extent do you agree with the following statements?	Strongly agree, n (%)	Agree, n (%)	Disagree, n (%)	Strongly disagree, n (%)
It was simple to use this app	90 (65.2)	46 (33.3)	1 (0.7)	1 (0.7)
It was easy to find the information I needed	51 (37)	68 (49.3)	13 (9.4)	6 (4.4)
It was easy to converse and type responses into this app	36 (26.1)	59 (42.8)	33 (23.9)	10 (7.3)
The information on the app screen is well-organized	59 (42.8)	76 (55.1)	1 (0.7)	2 (1.5)
It was easy to learn how to use the app	94 (68.1)	43 (31.2)	1 (0.7)	0 (0)
The symbols and buttons are easy to use	61 (44.2)	71 (51.5)	5 (3.6)	1 (0.7)
I understood how the app works the first time I used it	76 (58.1)	53 (38.4)	6 (4.4)	3 (2.2)
I can quickly remember how to use the app after a while	88 (63.8)	48 (34.8)	1 (0.7)	1 (0.7)
When I made a mistake using the app, I could easily correct the mistake	37 (26.8)	69 (50)	29 (21.1)	3 (2.2)
The app offered error messages that clearly told me how to fix the issues	22 (15.9)	57 (41.3)	52 (37.7)	7 (5.1)
The app could easily repeat words or statements that were not well understood	91 (66)	39 (28.3)	7 (5.1)	1 (0.7)
The app quickly loads the information on the screen	70 (50.7)	58 (42)	8 (5.8)	2 (1.5)
The health education content provided by the app was well-organized and delivered in a timely way	75 (54.3)	57 (41.3)	5 (3.6)	1 (0.7)
I was able to complete tasks quickly using the app	68 (49.3)	64 (46.4)	3 (2.2)	3 (2.2)
The app information was effective in helping me make decisions	74 (53.6)	60 (43.5)	1 (0.7)	3 (2.2)
The app has not stopped working or has ever closed	65 (47.1)	51 (37)	19 (13.8)	3 (2.2)
I believe the app could quickly improve health education and counseling	85 (61.59)	49 (35.51)	1 (0.72)	3 (2.1)
The app interface is nice to use	67 (48.6)	63 (45.7)	4 (2.9)	4 (2.9)
I enjoyed using the app interface	78 (56.5)	55 (39.9)	3 (2.2)	2 (1.5)
I am satisfied with the app and intend to continue using it	78 (56.5)	49 (35.5)	8 (5.8)	3 (2.2)
I want to encourage others to use the app	91 (66)	44 (31.9)	2 (1.5)	1 (0.7)
The app was culturally appropriate and I could easily share it with a family member or community member	79 (57.3)	54 (39.1)	5 (3.6)	0 (0)

Survey Responses by Age, Marital, and Education Status

Table 3 shows selected survey responses of the 138 participants who completed the survey questions by age, education, and marital status.

Participants aged 18-24 years (vs those aged 25-28 years) agreed that the "Nthabi app was simple to use" (106/106, 100% vs 30/32, 93.8%; *P*=.01), and agreed that "the educational content was well organised and delivered in a timely way" (104/106, 98.1% vs 28/32, 87.5%; *P*=.01).

Participants who were married (vs unmarried) agreed that "the educational content was well organised and delivered in a timely way" (9/11, 81.8% vs 123/127, 96.9%; *P*=.02), and agreed that "the app was nice to use" (9/11, 81.8% vs 121/127, 95.3%; *P*=.07).

Finally, young women who were in high school (vs those in tertiary education) were more likely to agree that "the app offered error messages that clearly told me how to fix the issue" (37/56, 66.1% vs 41/80, 51.3%; *P*=.02), and were "satisfied with the application and intended to continue using it" (55/56, 98.2% vs 70/80, 87.5%; *P*=.07).

Taken together, these results indicate that younger women, those in high school (and usually younger), and those unmarried (and usually younger) perceived Nthabi more positively.

XSL•FO
RenderX

Table 3. Survey responses of young women using Nthabi app by age, marital, and educational status.

	Agree, n (%)	Disagree, n (%)	*P* value
Opinions of young women and their marital status			
The health education content provided by the app was well-organized and delivered in a timely way			.02
Married (n=11	9 (81.8)	2 (18.2)	
Not married (n=127)	123 (96.9)	4 (3.2)	
The app interface is nice to use			.07
Married (n=11)	9 (81.8)	2 (18.1)	
Not married (n=127)	121 (95.3)	3 (2.4)	
Opinions of young women and their age range			
It was simple to use this app			.01
18-24 (n=106)	106 (100)	0 (0)	
25-28 (n=32)	30 (93.8)	2 (6.2)	
The health education content provided by the app was well organised and delivered in a timely way			.01
18-24 (n=106)	104 (98.1)	2 (1.9)	
25-28 (n=32)	28 (87.5)	4 (12.5)	
Opinions of young women and their educational status			
The app offered error messages that clearly told me how to fix the issues			.02
Primary school (n=2)	1 (50)	1 (50)	
High school (n=56)	37 (66.1)	19 (33.9)	
Tertiary (n=80)	41 (51.3)	39 (48.8)	
I am satisfied with the app and intend to continue using it			.07
Primary school (n=2)	2 (100)	0 (0)	
High school (n=56)	55 (98.2)	1 (1.8)	
Tertiary (n=80)	70 (87.5)	10 (12.5)	

Discussion

Principal Results

Young women in the lower middle-income country of Lesotho in southern Africa who used the newly adapted Nthabi intervention for up to 6 weeks perceived the usability and acceptability of the system very positively. Most respondents were satisfied with Nthabi and perceived it to be effective, efficient, and culturally appropriate. Participants agreed that Nthabi helped them make decisions and could improve the delivery of health education. They reported it was easy to use and well organized. Most intended to use it beyond the study period and they said they would encourage others to use it.

Improving sexual reproductive health education is a clear priority in Lesotho [9,10]. This study supports the idea that conversational agent technologies can provide sexual and reproductive health education in a rural, mountainous country like Lesotho, which has profound human resources challenges. As additional data are collected, the Ministry of Health and the health development and implementing partners should consider using Nthabi as a health promotion and education tool in Lesotho.

Comparison With Prior Work

These findings are in accordance with our previous research reporting results of focus groups of potential users who used an early version of Nthabi and key informant interviews of Ministry of Health officials. Participants reported that adaptations were culturally appropriate, and provided relevant and culturally sensitive clinical information. These qualitative data and now survey data together highlight the importance of acknowledging the local context when adapting an intervention. Nthabi was adapted to the uses, languages, interests, and realities of young people, as well as the importance of knowing what is preferred by young people as a measure of attractiveness to promote user engagement [24]. Most respondents were satisfied with the educational content and agreed that it delivered culturally appropriate and sensitive sexual and reproductive health information. Adaptations of interventions using appropriate cultural cues have a higher probability of acceptability and usability [25]. Culturally responsive interventions are effective in enhancing knowledge acquisition, attitudes, and satisfaction since they respect cultural diversity and the sociocultural factors that may affect health [26,27].

Participants agreed Nthabi could improve the delivery of health education and help them make health decisions. This finding is similar to findings from other studies conducted in lower

XSL•FO
RenderX

middle-income countries, which provide evidence that a variety of mHealth apps such as voice messages and daily educational text messages can improve young people's sexual reproductive health [28] and have been shown to be feasible and acceptable for improving health education and knowledge among adolescents and young people [29]. Other studies highlighted the broad potential for digital interventions to enhance health promotion and service delivery toward better sexual health [30,31]. However, this is the first study of the acceptability and usability of potentially more engaging and effective conversational agent systems in a low- and middle-income country in southern Africa.

Younger women in this study sample appear to have more positive perceptions of Nthabi than older participants. They found the system simple to use and the content delivered in a way convenient to them. Younger women might be more familiar and comfortable with using new technologies. This is consistent with other studies of women from the global north showing their preference for digital technologies such as readily available information, and their preference for opportunities to learn more about their bodies and health status [32]. Other studies have found that younger people are not only accepting of new technologies in health care settings but are actually looking for more of these technologies to use in health settings [33,34].

Accessibility of Nthabi on Mobile Devices

In Lesotho, 94% of people aged 18-29 years use smartphones, and 3G data coverage is available in almost 90% of the country [5], yet access to public Wi-Fi and data costs remain barriers to using mobile technologies for health education. Nthabi was designed to address our concern that limited internet access would impact participants' use of Nthabi. A decision was made to download the full system to mobile devices so that participants could use the system when not in Wi-Fi environments. While this design allowed the participants to use the system at their convenience, the inclusion of all the content and most importantly, the inclusion of the system voice synthesizer, created significant difficulties for downloading and using Nthabi on most phones due to low phone memory. The finding that only 8 of 172 (4.7%) participants were able to use Nthabi on their phones demonstrates that mobile phone use is possible, though practically, only phones with sufficient available memory could be used. As it becomes increasingly possible for young women to have regular access to public Wi-Fi, it will become possible for more young women to use Nthabi in the cloud on their phones rather than downloading the full system.

Participants who were unable to download the intervention to their phones were loaned tablet devices. We purchased 20 devices (US $111 per tablet or US $14 per participant) and loaned them to participants on a rolling basis. At the end of the study, all tablets were returned. While this is a cost-effective alternative, future studies of large-scale health education programs in low-resource settings using cloud-based interventions will be possible with increased public Wi-Fi availability. We are now planning studies in which fully downloadable and Wi-Fi–enabled systems are available.

Limitations

There are several limitations to this study. First, the results are not nationally representative of women from Lesotho as participants were recruited by convenience from only 2 of the 10 districts of Lesotho. The sample included many participants recruited from the university and vocational schools, and while these participants reported residing in and receiving health services in Berea and Leribe, the results do not necessarily reflect the views of women living in rural areas. Further trials are needed to more definitively identify the perceptions of rural women. A larger study in all 10 districts of Lesotho is planned.

Second, while this study reports on perceptions of successful usability and acceptability, it does not provide evidence that the intervention improved young women's health knowledge, attitudes, and behaviors. Research to further determine the impact of knowledge of the topics discussed by Nthabi is underway.

Conclusions

Nthabi is a potentially useful intervention for providing sexual reproductive health information for young women in the rural, lower middle-income country of Lesotho with limited human resources in health. Further study of the Nthabi system is warranted to determine if the Nthabi health education content and interactive dialogue about sexual and reproductive health can improve women's knowledge, attitudes, and health behaviors.

Acknowledgments

All authors made substantial contributions to draft and critically revise the manuscript. EN-N, MM-M, and BWJ contributed to conceptualizing the study and to data acquisition, analysis, interpretation, and writing the manuscript. TB contributed to conceptualizing the study and as an expert in conversational agents. CJ contributed study design, data collection, and manuscript writing. All authors approved the final draft and agreed to be accountable for all aspects of the work. The authors are grateful to the Lesotho Ministry of Health leadership, young women who participated in this study, and research assistants Paballo Lethunya and Moleboheng Mofolo. The research reported in this publication was supported by the Fogarty International Center of the United States National Institutes of Health (award R21TW011361). The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

Conflicts of Interest

None declared.

XSL•FO
RenderX

Multimedia Appendix 1
Brief Description of Nthabi Adaptation.
[DOCX File , 16 KB - humanfactors_v11i1e52048_app1.docx]

References

1. Kabir MF, Schulman D, Abdullah AS. Promoting relational agent for health behavior change in low and middle - income counntries (LMICs): issues and approaches. J Med Syst 2019;43(7):227 [FREE Full text] [doi: 10.1007/s10916-019-1360-z] [Medline: 31190131]
2. Balaji D, He L, Giani S, Bosse T, Wiers R, de Bruijn GJ. Effectiveness and acceptability of conversational agents for sexual health promotion: a systematic review and meta-analysis. Sex Health 2022;19(5):391-405 [FREE Full text] [doi: 10.1071/SH22016] [Medline: 35863761]
3. Bacchus LJ, Reiss K, Church K, Colombini M, Pearson E, Naved R, et al. Using digital technology for sexual and reproductive health: are programs adequately considering risk? Glob Health Sci Pract 2019;7(4):507-514 [FREE Full text] [doi: 10.9745/GHSP-D-19-00239] [Medline: 31874936]
4. Fernández C, Vicente MA, Carrillo I, Guilabert M, Mira JJ. Factors influencing the smartphone usage behavior of pedestrians: observational study on "Spanish smombies". J Med Internet Res 2020;22(8):e19350 [FREE Full text] [doi: 10.2196/19350] [Medline: 32667896]
5. Pule N. The state of ICT in Lesotho (2016). Commissioned by the Lesotho Communications Authority and the International Telecommunication Union. Research ICT Africa. 2017. URL: https://researchictafrica.net/wp/wp-content/uploads/2018/01/2017_The-State-of-ICT-in-Lesotho_RIA_LCA.pdf [accessed 2023-12-04]
6. Naidoo K, Adeagbo O, Pleaner M. Sexual and reproductive health needs of adolescent girls and young women in sub-Saharan Africa: research, policy, and practice. SAGE Open 2019;9(3):2158244019859951 [FREE Full text] [doi: 10.1177/2158244019859951]
7. Mwamba B, Mayers P, Shea J. Sexual and reproductive health knowledge of postgraduate students at the University of Cape Town, in South Africa. Reprod Health 2022;19(1):225 [FREE Full text] [doi: 10.1186/s12978-022-01507-6] [Medline: 36522681]
8. Rokicki S, Fink G. Assessing the reach and effectiveness of mHealth: evidence from a reproductive health program for adolescent girls in Ghana. BMC Public Health 2017;17(1):969 [FREE Full text] [doi: 10.1186/s12889-017-4939-7] [Medline: 29262823]
9. Lesotho population-based HIV impact assessment 2020. (LePHIA 2020). Lesotho Ministry of Health. 2020. URL: https://phia.icap.columbia.edu/countries/lesotho/ [accessed 2023-12-04]
10. Lesotho's number of maternal deaths 1990–2015. World Bank. URL: https://www.ceicdata.com/en/lesotho/health-statistics/ls-number-of-maternal-death [accessed 2023-12-04]
11. World Bank. 2018. URL: https://www.worldbank.org/en/topic/health/brief/lesothos-health-sector [accessed 2023-12-04]
12. Bickmore TW, Pfeifer LM, Byron D, Forsythe S, Henault LE, Jack BW, et al. Usability of conversational agents by patients with inadequate health literacy: evidence from two clinical trials. J Health Commun 2010;15(Suppl 2):197-210 [FREE Full text] [doi: 10.1080/10810730.2010.499991] [Medline: 20845204]
13. Provoost S, Lau HM, Ruwaard J, Riper H. Embodied conversational agents in clinical psychology: a scoping review. J Med Internet Res 2017;19(5):e151 [FREE Full text] [doi: 10.2196/jmir.6553] [Medline: 28487267]
14. Abd-Alrazaq AA, Alajlani M, Ali N, Denecke K, Bewick BM, Househ M. Perceptions and opinions of patients about mental health chatbots: scoping review. J Med Internet Res 2021;23(1):e17828 [FREE Full text] [doi: 10.2196/17828] [Medline: 33439133]
15. Jack B, Bickmore T, Hempstead M, Yinusa-Nyahkoon L, Sadikova E, Mitchell S, et al. Reducing preconception risks among African American women with conversational agent technology. J Am Board Fam Med 2015;28(4):441-451 [FREE Full text] [doi: 10.3122/jabfm.2015.04.140327] [Medline: 26152434]
16. Jack BW, Bickmore T, Yinusa-Nyahkoon L, Reichert M, Julce C, Sidduri N, et al. Improving the health of young African American women in the preconception period using health information technology: a randomised controlled trial. Lancet Digit Health 2020;2(9):e475-e485 [FREE Full text] [doi: 10.1016/S2589-7500(20)30189-8] [Medline: 33328115]
17. Nkabane-Nkholongo E, Mokgatle M, Bickmore T, Julce C, Jack BW. Adaptation of the Gabby conversational agent system to improve the sexual and reproductive health of young women in Lesotho. Front Digit Health 2023;5:1224429 [FREE Full text] [doi: 10.3389/fdgth.2023.1224429] [Medline: 37860039]
18. Manzano-Monfort G, Paluzie G, Díaz-Gegúndez M, Chabrera C. Usability of a mobile application for health professionals in home care services: a user-centered approach. Sci Rep 2023;13(1):2607 [FREE Full text] [doi: 10.1038/s41598-023-29640-7] [Medline: 36788261]
19. Hajesmaeel-Gohari S, Khordastan F, Fatehi F, Samzadeh H, Bahaadinbeigy K. The most used questionnaires for evaluating satisfaction, usability, acceptance, and quality outcomes of mobile health. BMC Med Inform Decis Mak 2022;22(1):22 [FREE Full text] [doi: 10.1186/s12911-022-01764-2] [Medline: 35081953]

XSL•FO
RenderX

20. Gardiner P, Hempstead MB, Ring L, Bickmore T, Yinusa-Nyahkoon L, Tran H, et al. Reaching women through health information technology: the Gabby preconception care system. Am J Health Promot 2013;27(3 Suppl):eS11-eS20 [FREE Full text] [doi: 10.4278/ajhp.1200113-QUAN-18] [Medline: 23286652]
21. Walker R, Drakeley S, Julce C, Sidduri N, Bickmore T, Skouteris H, et al. Assessing the potential of a virtual patient advocate to provide preconception care and health advice to women living in Australia. J Dev Orig Health Dis 2021;12(2):300-308 [FREE Full text] [doi: 10.1017/S2040174420000264] [Medline: 32308183]
22. Schulz KF, Altman DG, Moher D, CONSORT Group. CONSORT 2010 statement: updated guidelines for reporting parallel group randomised trials. BMJ 2010;340:c332 [FREE Full text] [doi: 10.1136/bmj.c332] [Medline: 20332509]
23. Eysenbach G, CONSORT-EHEALTH Group. CONSORT-EHEALTH: improving and standardizing evaluation reports of web-based and mobile health interventions. J Med Internet Res 2011;13(4):e126 [FREE Full text] [doi: 10.2196/jmir.1923] [Medline: 22209829]
24. Moeini B, Rezapur-Shahkolai F, Jahanfar S, Naghdi A, Karami M, Ezzati-Rastegar K. Utilizing the PEN-3 model to identify socio-cultural factors affecting intimate partner violence against pregnant women in suburban Hamadan. Health Care Women Int 2019;40(11):1212-1228 [FREE Full text] [doi: 10.1080/07399332.2019.1578777] [Medline: 31166150]
25. Arruzza E, Chau M. The effectiveness of cultural competence education in enhancing knowledge acquisition, performance, attitudes, and student satisfaction among undergraduate health science students: a scoping review. J Educ Eval Health Prof 2021;18:3 [FREE Full text] [doi: 10.3352/jeehp.2021.18.3] [Medline: 33621460]
26. Echeverri M, Chen AMH. Educational interventions for culturally competent healthcare: developing a protocol to conduct a systematic review of the rationale, content, teaching methods, and measures of effectiveness. J Best Pract Health Prof Divers 2016;9(1):1160-1177 [FREE Full text] [Medline: 34296221]
27. Martin P, Cousin L, Gottot S, Bourmaud A, de La Rochebrochard E, Alberti C. Participatory interventions for sexual health promotion for adolescents and young adults on the internet: systematic review. J Med Internet Res 2020;22(7):e15378 [FREE Full text] [doi: 10.2196/15378] [Medline: 32735217]
28. Feroz AS, Ali NA, Khoja A, Asad A, Saleem S. Using mobile phones to improve young people sexual and reproductive health in low and middle-income countries: a systematic review to identify barriers, facilitators, and range of mHealth solutions. Reprod Health 2021;18(1):9 [FREE Full text] [doi: 10.1186/s12978-020-01059-7] [Medline: 33453723]
29. Mitchell KJ, Bull S, Kiwanuka J, Ybarra ML. Cell phone usage among adolescents in Uganda: acceptability for relaying health information. Health Educ Res 2011;26(5):770-781 [FREE Full text] [doi: 10.1093/her/cyr022] [Medline: 21536715]
30. Sewak A, Yousef M, Deshpande S, Seydel T, Hashemi N. The effectiveness of digital sexual health interventions for young adults: a systematic literature review (2010-2020). Health Promot Int 2023;38(1):daac104 [FREE Full text] [doi: 10.1093/heapro/daac104] [Medline: 36757346]
31. Kouvari M, Karipidou M, Tsiampalis T, Mamalaki E, Poulimeneas D, Bathrellou E, et al. Digital health interventions for weight management in children and adolescents: systematic review and meta-analysis. J Med Internet Res 2022;24(2):e30675 [FREE Full text] [doi: 10.2196/30675] [Medline: 35156934]
32. Lupton D. Young people's use of digital health technologies in the Global North: narrative review. J Med Internet Res 2021;23(1):e18286 [FREE Full text] [doi: 10.2196/18286] [Medline: 33427684]
33. Dietrich JJ, Coetzee J, Otwombe K, Mdanda S, Nkala B, Makongoza M, et al. Adolescent-friendly technologies as potential adjuncts for health promotion. Health Education 2014;114(4):304-318 [FREE Full text] [doi: 10.1108/he-08-2013-0039]
34. Radovic A, McCarty CA, Katzman K, Richardson LP. Adolescents' perspectives on using technology for health: qualitative study. JMIR Pediatr Parent 2018;1(1):e2 [FREE Full text] [doi: 10.2196/pediatrics.8677] [Medline: 30740590]

Abbreviations

CONSORT: Consolidated Standards of Reporting Trials

Edited by A Choudhury; submitted 21.08.23; peer-reviewed by M Brown, J Xie; comments to author 07.11.23; revised version received 08.12.23; accepted 11.12.23; published 12.03.24.

Please cite as:
Nkabane-Nkholongo E, Mpata-Mokgatle M, Jack BW, Julce C, Bickmore T
Usability and Acceptability of a Conversational Agent Health Education App (Nthabi) for Young Women in Lesotho: Quantitative Study
JMIR Hum Factors 2024;11:e52048
URL: https://humanfactors.jmir.org/2024/1/e52048
doi:10.2196/52048
PMID:38470460

XSL•FO
RenderX

©Elizabeth Nkabane-Nkholongo, Mathildah Mpata-Mokgatle, Brian W Jack, Clevanne Julce, Timothy Bickmore. Originally published in JMIR Human Factors (https://humanfactors.jmir.org), 12.03.2024. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on https://humanfactors.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Stimulating Preconception Care Uptake by Women With a Vulnerable Health Status Through a Mobile Health App (Pregnant Faster): Pilot Feasibility Study

Sharissa M Smith[1], MD; Babette Bais[1], PhD; Hafez Ismaili M'hamdi[2], PhD; Maartje HN Schermer[2], Prof Dr; Régine PM Steegers-Theunissen[1], Prof Dr, MD

[1]Department of Obstetrics and Gynecology, Erasmus University Medical Center, Rotterdam, Netherlands

[2]Department of Medical Ethics, Philosophy and History of Medicine, Erasmus University Medical Center, Rotterdam, Netherlands

Corresponding Author:
Régine PM Steegers-Theunissen, Prof Dr, MD
Department of Obstetrics and Gynecology
Erasmus University Medical Center
Doctor Molewaterplein 40
Rotterdam, 3015 CD
Netherlands
Phone: 31 10704 ext 0704
Email: r.steegers@erasmusmc.nl

Abstract

Background: A low socioeconomic status is associated with a vulnerable health status (VHS) through the accumulation of health-related risk factors, such as poor lifestyle behaviors (eg, inadequate nutrition, chronic stress, and impaired health literacy). For pregnant women, a VHS translates into a high incidence of adverse pregnancy outcomes and therefore pregnancy-related inequity. We hypothesize that stimulating adequate pregnancy preparation, targeting lifestyle behaviors and preconception care (PCC) uptake, can reduce these inequities and improve the pregnancy outcomes of women with a VHS. A nudge is a behavioral intervention aimed at making healthy choices easier and more attractive and may therefore be a feasible way to stimulate engagement in pregnancy preparation and PCC uptake, especially in women with a VHS. To support adequate pregnancy preparation, we designed a mobile health (mHealth) app, Pregnant Faster, that fits the preferences of women with a VHS and uses nudging to encourage PCC consultation visits and engagement in education on healthy lifestyle behaviors.

Objective: This study aimed to test the feasibility of Pregnant Faster by determining usability and user satisfaction, the number of visited PCC consultations, and the course of practical study conduction.

Methods: Women aged 18-45 years, with low-to-intermediate educational attainment, who were trying to become pregnant within 12 months were included in this open cohort. Recruitment took place through social media, health care professionals, and distribution of flyers and posters from September 2021 until June 2022. Participants used Pregnant Faster daily for 4 weeks, earning coins by reading blogs on pregnancy preparation, filling out a daily questionnaire on healthy lifestyle choices, and registering for a PCC consultation with a midwife. Earned coins could be spent on rewards, such as fruit, mascara, and baby products. Evaluation took place through the mHealth App Usability Questionnaire (MAUQ), an additional interview or questionnaire, and assessment of overall study conduction.

Results: Due to limited inclusions, the inclusion criterion “living in a deprived neighborhood” was dropped. This resulted in the inclusion of 47 women, of whom 39 (83%) completed the intervention. In total, 16 (41%) of 39 participants visited a PCC consultation, with their main motivation being obtaining personalized information. The majority of participants agreed with 16 (88.9%) of 18 statements of the MAUQ, indicating high user satisfaction. The mean rating was 7.7 (SD 1.0) out of 10. Points of improvement included recruitment of the target group, simplification of the log-in system, and automation of manual tasks.

Conclusions: Nudging women through Pregnant Faster to stimulate pregnancy preparation and PCC uptake has proven feasible, but the inclusion criteria must be revised. A substantial number of PCC consultations were conducted, and this study will therefore be continued with an open cohort of 400 women, aiming to establish the (cost-)effectiveness of an updated version, named Pregnant Faster 2.

International Registered Report Identifier (IRRID): RR2-10.2196/45293

XSL•FO
RenderX

(JMIR Hum Factors 2024;11:e53614) doi:10.2196/53614

KEYWORDS

preconception care; mHealth; mobile health; pregnancy preparation; nudge; health inequality; socioeconomic status; lifestyle; women; pregnancy; pregnant women; pregnant; socioeconomic; pilot feasibility study; mHealth app; mHealth application; app; application; risk factor; nutrition; stress; chronic stress; health literacy; usability; user satisfaction; user; users

Introduction

A low socioeconomic status (SES) is associated with a vulnerable health status (VHS), which research suggests is grounded in the accumulation of risk factors, such as inadequate nutrition, smoking, and increased mental stressors [1-4]. For women, a low SES means they are more likely to have a VHS, which translates into a higher incidence of adverse pregnancy outcomes in this group [5-9]. These adverse outcomes originate at least partly in the periconception period [10], during which gametogenesis, embryonic development, and placentation take place, laying the foundation for perinatal outcomes, as well as the child's lifelong health [11,12]. For example, an accumulation of 2 or more maternal risk factors impacts embryonic growth [13], which is associated with midpregnancy fetal weight and birth weight [14]. In addition, infants born small for their gestational age are more susceptible to noncommunicable diseases, such as diabetes mellitus and cardiovascular disease [15]. The effects of adverse pregnancy outcomes therefore hit twice: once in utero and once in later life. This increases the child's chance of a VHS in adulthood, which, once again, may influence pregnancy outcomes. These transgenerational effects are further maintained by impaired health agency, which is associated with a low SES and diminishes the likelihood of seeking necessary care [16]. In accordance with these findings, research shows that women with a VHS are less likely to engage in pregnancy preparation and take up preconception care (PCC) [17].

PCC is usually given by a midwife or an obstetrician and is aimed at identification of possible risk factors for adverse outcomes, ameliorating those that are modifiable prior to pregnancy [18]. This includes adopting healthy lifestyle behaviors and making beneficial choices in general that will increase the chance of having a healthy pregnancy and baby. Although ≥80% of women who wish to become pregnant have at least 1 modifiable risk factor for adverse pregnancy outcomes [19,20], the uptake of PCC remains low due to insufficient awareness of risk factors and the benefits of PCC [21]. Women with a low SES may encounter additional barriers when engaging in pregnancy preparation, as they are already burdened by the deprived circumstances in which they live. Supporting this group by making pregnancy preparation easier and attractive might be a suitable way to relieve the inequity regarding their pregnancy outcomes.

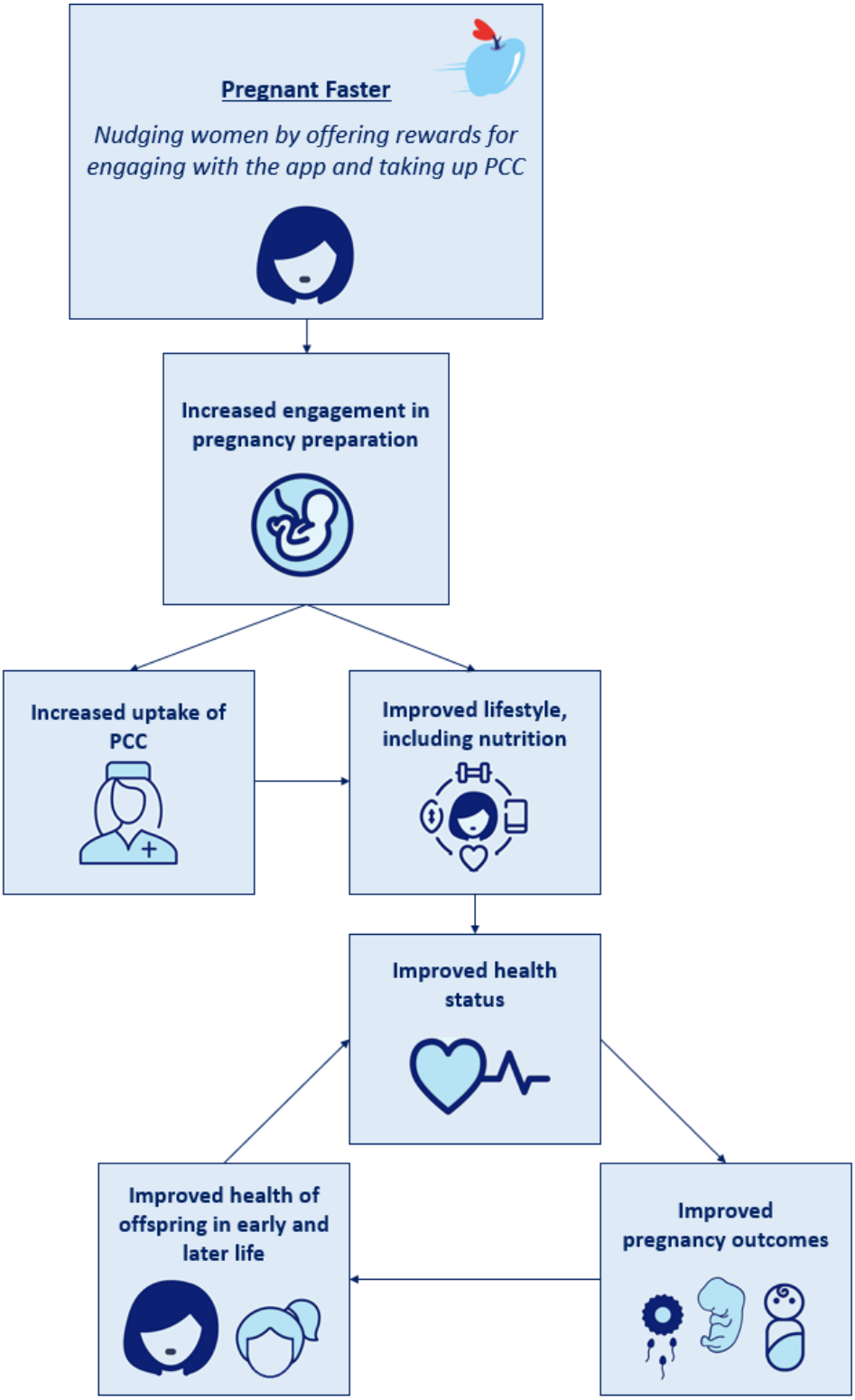
Our research group has previously developed the web-based PCC tool Smarter Pregnancy, an interactive, tailored, mobile health (mHealth) platform that offers practical coaching and customized feedback on nutrition and other lifestyle behaviors of prospective parents [22]. Smarter Pregnancy has proven to be effective in supporting healthy choices in women with a VHS, in addition to being valued highly by them [23]. To further support PCC engagement in women with a VHS, we have designed an mHealth app that especially fits their needs and preferences [24]: the app-based nudge Pregnant Faster. A nudge is an intervention that stimulates making beneficial choices by increasing the attractiveness and easiness of healthy behavior [25]. An in-depth explanation of nudge theory and its application in health policy can be found in the study by Murayama et al [26].

In the case of Pregnant Faster, participants are nudged through a loyalty program that entails collecting coins by engaging with the app and ordering rewards using those coins. The design of Pregnant Faster can be viewed as a macrolevel nudge, containing multiple microlevel nudges aimed at stimulating pregnancy preparation and encouraging the uptake of PCC. For example, the monetary value of a coin is a microlevel nudge; it varies from €0.06 to €0.26 (US $0.06-$0.26), depending on the type of reward. Healthy rewards, such as folic acid supplements, are relatively cheap, steering participants toward picking them over luxury goods, while maintaining their freedom to choose. The most important feature of the app, which also yields the highest number of coins, is the possibility to register for a PCC consultation with a nearby midwife, promoting blended care: an effective way to promote pregnancy preparation [27]. As midwives are the primary health care providers for pregnant women in the Netherlands, PCC consultations are beneficial for the bond between health care provider and client prior to and during pregnancy. The full description of Pregnant Faster's design process, detailing the imbedded nudges, has been published in *JMIR Protocols* [28].

The aim of this pilot study was to determine Pregnant Faster's feasibility pertaining to usability and user satisfaction, the number of PCC consultations booked and visited by participants, and the course of practical conduction regarding the inclusion process, reward allocation, and finalization of the study. In addition, the results of this study will be used to further develop Pregnant Faster and lay the foundation for a larger cohort study to establish its (cost-)effectiveness. Our overall ambition is that Pregnant Faster contributes to the improvement of short-term and long-term health in mothers with a VHS, their children, and future generations (Figure 1).

XSL•FO
RenderX

Figure 1. Aim of the mHealth app Pregnant Faster. mHealth: mobile health; PCC: preconception care.



Methods

Ethical Considerations

This study was assessed and approved by the Medical Ethical Committee of the Erasmus University Medical Center, Rotterdam, The Netherlands (MEC-2020-0974). Informed consent was obtained from all participants via email. Considering the low risk of this study, composing a Data Safety Monitoring Board was deemed unnecessary.

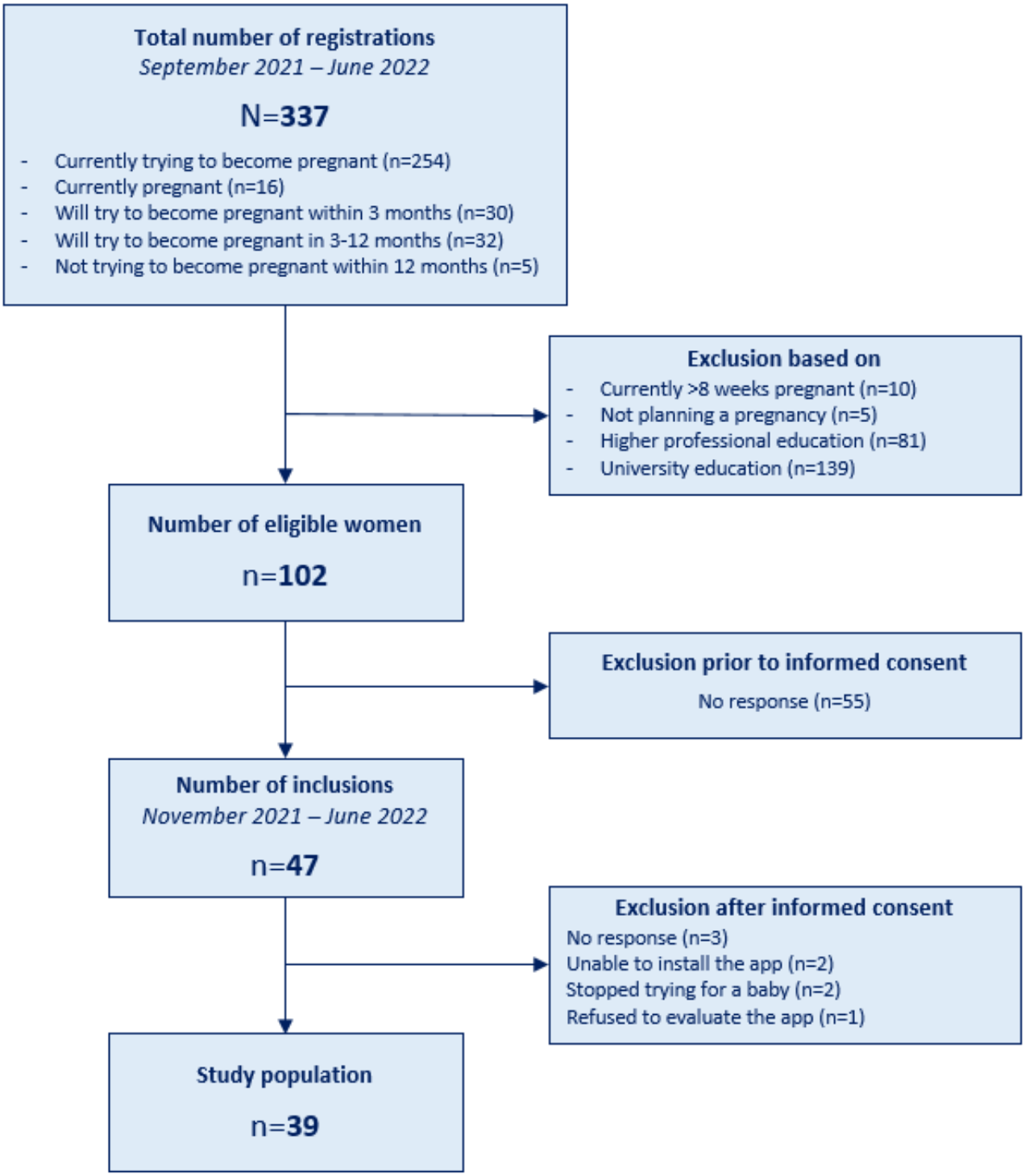
Recruitment and Inclusion

Our aim was to include 40 participants in this study. Between September 2021 and June 2022, 337 women registered for this study, of which 102 (30.3%) were eligible for inclusion. Due to a higher-than-expected confidentiality, integrity, and availability (CIA) Triad classification (a risk score regarding

XSL•FO
RenderX

user information safety [29]), additional security demands were necessary. Fulfilling these demands delayed the launch of Pregnant Faster from September 2021 to November 2021, leading to a loss of recruited eligible women. From the launch onward, 47 (46.1%) women were included, of which 39 (83%) completed the intervention (Figure 2).

Figure 2. Inclusion flowchart.



Recruitment took place through the Sneller Zwanger website [30], which was distributed through posters and flyers, advertisements on the social media platforms Facebook and Instagram, and midwifery practices that provide primary care to all pregnant women in the Netherlands. Additionally, a collaboration took place with the Dutch influencer Midwife Mother (Dutch *Verlosmoeder*) on Instagram [31-33]. Participants filled in a survey with their first name, age, telephone number, email address, zip code, educational level, and when they planned on trying to become pregnant (currently pregnant, currently trying, or trying in ≤3, >3-12, or >12 months). Pregnant women were asked to fill in their estimated or calculated due date.

Participants were selected based on the following inclusion criteria: assigned female at birth, 18-45 years old, actively trying

XSL•FO
RenderX

to become pregnant within now and 12 months or pregnant with a gestational age of <8 weeks at the start of the intervention, a low-to-intermediate educational level (prevocational or vocational education), and able and willing to download and evaluate the app.

Exclusion criteria were as follows: insufficient proficiency in the Dutch language, not in possession of a smartphone or tablet suitable for the app, and refusal to download or evaluate the app. All excluded women received a free coupon for the Dutch or English version of Smarter Pregnancy [34].

Design

Pregnant Faster was developed by the Erasmus University Medical Center's research group Periconception Epidemiology at the Department of Obstetrics and Gynecology, in collaboration with TJIP The Platform Engineer and the event bureau Improve. A detailed description of the cocreation and design process of Pregnant Faster and the study protocol has been published in *JMIR Protocols* [28].

Intervention

Eligible women were sent the patient information folder, in which the intervention was explained. Inclusion was finalized after a telephone conversation in which further clarification could be provided. An email was sent with instructions on how to download and install Pregnant Faster from Apple App Store (iOS) or via a link (Android). If more than 3 days passed between inclusion and downloading, participants were approached twice by email and telephone and once by a text message to provide further support with installation.

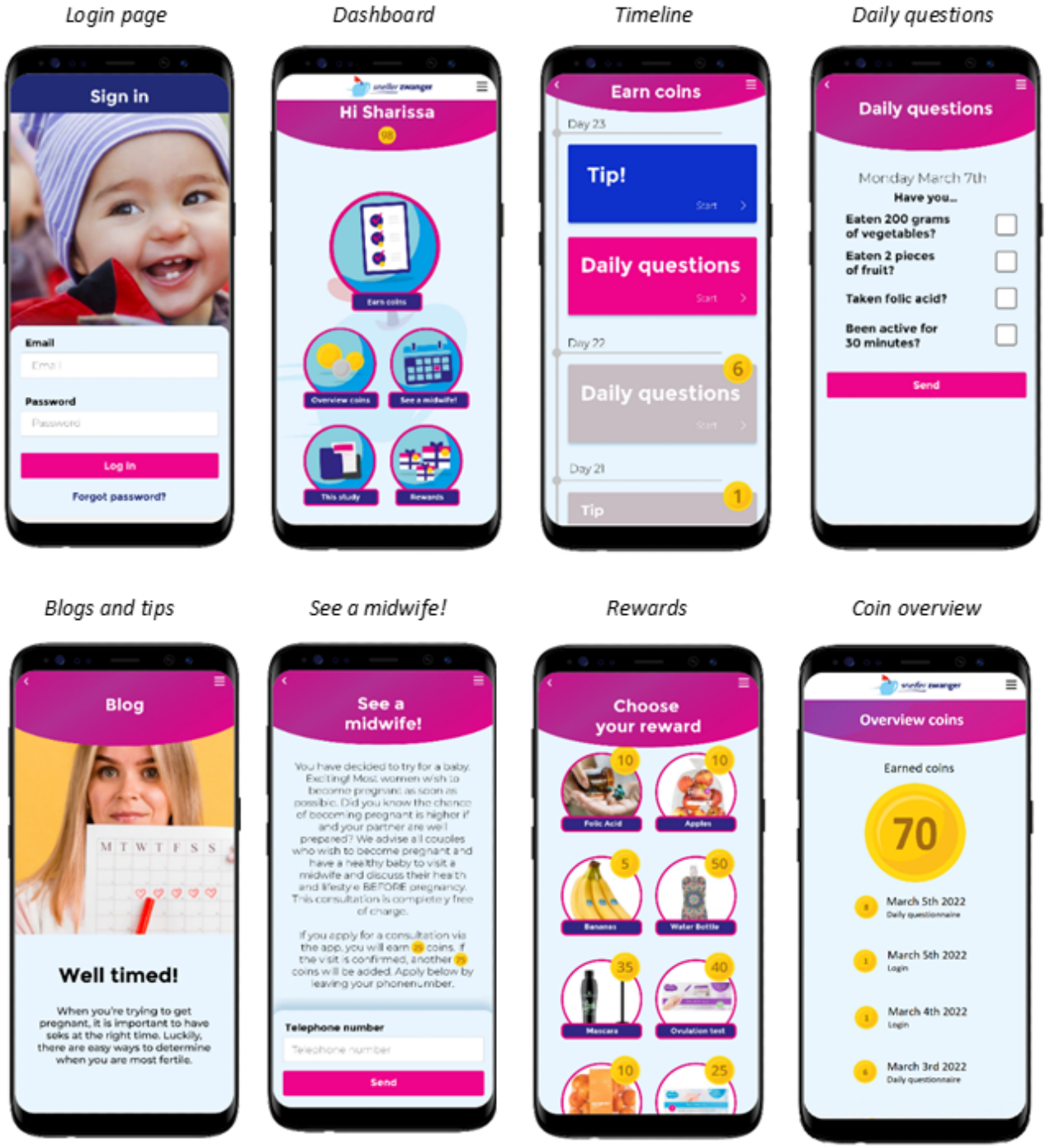
The first log-in marked the start of the 4-week intervention. During this period, participants logged in with their email address and a password, which yielded 1 coin per day per log-in. The first log-in yielded 50 coins as a reward for installation and to immediately stimulate participants to further engage with the app. After log-in, a dashboard appeared, containing 5 buttons: (1) "Earn coins," (2) "Overview coins," (3) "See a midwife!," (4) "This study," and (5) "Rewards" (Figure 3).

Button 1, "Earn coins," led to a timeline where new blogs and tips appeared daily (Multimedia Appendix 1). Reading this information yielded 4-8 coins. In the same timeline, a daily questionnaire appeared in which participants could tick a box if they ate sufficient fruit and vegetables, exercised, and took folic acid supplements that day. Each ticked box yielded 2 coins per day. Button 2, "Overview coins," displayed when and how coins were earned and how many coins were spent on which products. Button 3, "See a midwife!," contained information regarding what PCC is and who it is for, stimulating participants to register through the app for a PCC consultation. Registering consisted of filling in their phone number, which immediately yielded 25 coins. An additional 75 coins were allocated after the visit was confirmed by the midwife. Button 4, "This study," contained information about the study itself and contact details for support. Button 5, "Rewards," contained an in-app shop where participants could order rewards, including (but not limited to) folic acid supplements, fruit, nail polish, mascara, ovulation or pregnancy tests, and newborn clothing. Rewards were sent to their home address to arrive within 5 business days. If a participant had not logged in for 7 days, they received a manually sent text message and email, encouraging them to read up on the newly offered blogs and tips, earn more coins, and order rewards.

At the end of the intervention, participants were offered a coupon for Smarter Pregnancy via the timeline, which would yield 25 coins upon use and provide them with an additional 26 weeks of coaching. Furthermore, they received an email regarding finalization of the study and available means of support. Earned coins could be spent up to 2 weeks after the intervention ended. The blogs and tips remained accessible for as long as the app remained installed. Figure 4 provides an overview of the study flow.

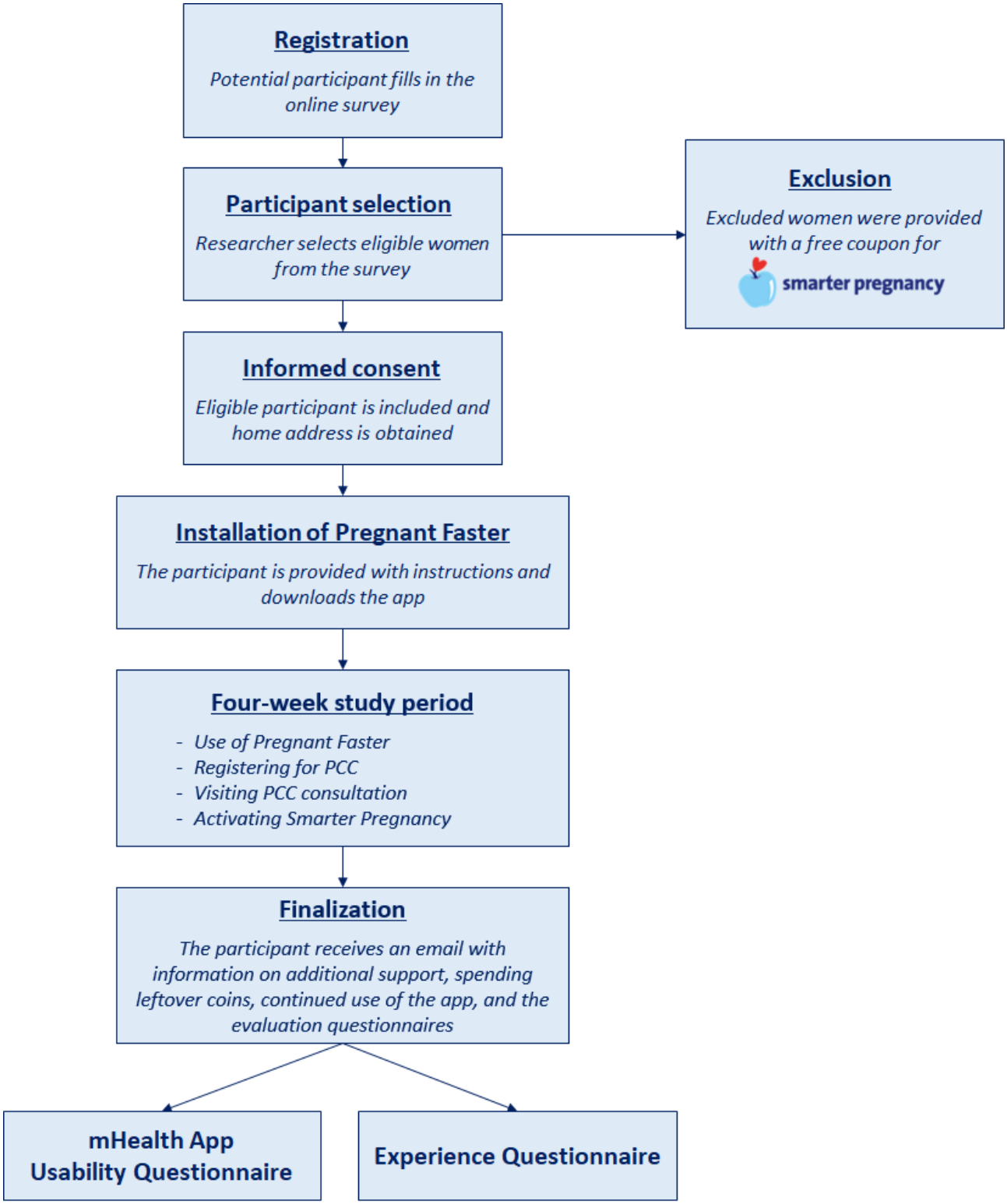
XSL•FO
RenderX

Figure 3. Pregnant Faster interface.



XSL•FO
RenderX

Figure 4. Flowchart of the study design. mHealth: mobile health; PCC: preconception care.



Data Collection and Outcome Measures

During the study, registration and inclusion rates were tracked and a log book was kept to note any encountered barriers and changes to the protocol.

Participants' baseline characteristics were collected prior to inclusion through the selection survey, and home addresses were collected via email after obtaining informed consent. If selected participants did not respond to attempts to include them, they were contacted twice by email, twice by telephone, and once by a text message.

After using Pregnant Faster for 4 weeks, participants filled in a modified version of the 18-item mHealth App Usability Questionnaire (MAUQ), which uses a 7-point Likert scale (Multimedia Appendix 2) [35]. In addition, the first 10 (25.6%) participants went through a semistructured interview (Multimedia Appendix 3) in which they elaborated on their experiences with the app and, if applicable, the PCC

XSL•FO
RenderX

consultation. The audiotapes of these interviews were used to compose the Experience Questionnaire (ExQ; Multimedia Appendix 4) that consisted partly of questions using a 5-point Likert scale. The ExQ was offered to the remaining participants, and the first author filled in the ExQ for the first 10 (25.6%) participants, using the audiotaped data provided during the interviews. If a question in the ExQ was not clearly answered in the interview, the participant was approached by telephone to provide an answer. The answers to the open questions in the ExQ were evaluated for notable and recurring comments.

At the end of the study period, data were collected on the number of coins earned, the types of rewards that were chosen, and the number of booked and visited PCC consultations.

Data Analysis

To evaluate the inclusion strategy, the following percentages were calculated: (1) eligibility percentage, (2) inclusion percentages, and (3) intervention completion percentages. The eligibility percentage was determined by dividing the number of women eligible for inclusion by the total number of women who registered for the study. The inclusion percentages were calculated by dividing the number of included participants (who provided informed consent) by the number of total registrations and the number of eligible women who were approached for inclusion. Completion of the intervention entailed completing the evaluation of the app. The intervention completion percentages were obtained by dividing the number of participants who completed the intervention by the number of eligible women and the number of participants who provided informed consent.

The baseline of the study population is presented in tabular form using the median (IQR) for continuous data and n (%) for categorical data. Data obtained through the MAUQ and the ExQ are presented in bar charts. Notable answers to open questions are presented in narrative form. Data on feasibility from the researchers' point of view are presented as bullet points.

All calculations were carried out using SPSS Statistics 25 (IBM Corporation), charts were created using Excel 2016 (Microsoft), and figures were created in PowerPoint 2016 (Microsoft).

Results

Recruitment, Inclusion, and the Study Population

A total of 337 women registered for the intervention, of whom 102 (30.3%) were eligible for inclusion. Informed consent was signed by 47 (46.1%) women, and 39 (83%) of the 47 women participated and completed the intervention. Table 1 displays participants' baseline characteristics.

Table 1. Participants' baseline characteristics.

Characteristics	Participants (N=39)
Age (years), median (IQR)	30 (27-35)
Mean income neighborhood[a], n (%)	
Below middle	19 (48.7)
Above middle	11 (28.2)
Low to high	4 (10.3)
High	5 (12.8)
Educational level[b], n (%)	
Low	3 (7.7)
Intermediate	36 (92.3)
Trying to become pregnant, n (%)	
Currently trying	32 (82.1)
Within 3 months	5 (12.8)
Within 12 months	2 (5.1)
Mobile operating system, n (%)	
Android	17 (43.6)
iOS	22 (56.4)

[a]The median household income of a neighborhood is determined by the distribution of household income of all households in the country [28]. This table adheres to the original subdivision of the distribution of household income (year 2020): low, <€15,900 (<US $18,800); below middle, €15,900-21,000 (US $18,800-$24,800); middle, €21,000-26,800 (US $24,800-$31,700); above middle, €26,800-34,600 (US $31,700-$40,900); and high, >€34,600 (>US $40,900).

[b]Educational level [29]. Dutch educational levels are subdivided as follows: low (prevocational education, selective secondary education, or lower), intermediate (vocational education), and high (bachelor's degree, master's degree, or higher).

A 2-month gap arose between the start of recruitment and the intervention, due to the app's CIA Triad classification [29]. A low classification was expected, but the combination of 40 intended participants and their registering their first name and

XSL•FO
RenderX

email address in the app warranted a slightly higher classification for confidentiality and therefore additional security demands. Despite frequent updates to keep eligible women engaged, 55 (53.9%) of 102 women did not respond when inclusion commenced. Next, we describe these events and the inclusion process in detail.

Between September 2021 and January 2022, 212 (62.9%) women from the total 337 registrations reached in June 2022 registered for the study. Of these 212 women, only 9 (4.2%) were included. To boost the registration and inclusion rates, more flyers and posters were distributed, and the choice was made to include women who were trying to become pregnant within 12 months as opposed to within 3 months, as originally intended. Furthermore, the intervention was expanded from the municipality of Rotterdam to nationwide, delivering rewards through the postal service instead of by car. Midwives throughout the Netherlands were actively approached to ask whether they were interested in participating in the study and were offered support in setting up PCC consultations in their practices. Subsequently, the collaboration with Midwife Mother was renewed, who uploaded another post and multiple stories regarding PCC and Pregnant Faster to her Instagram. All women who were previously excluded based on not living near Rotterdam were contacted and asked to participate.

These efforts showed a limited effect. By May 2022, an additional 54 (16%) registrations and a total of 16 (34%) inclusions were obtained, which led to the decision to drop the inclusion criterion of living in a deprived neighborhood, thereby lowering the chance of including women likely to have a VHS. This choice was made to allow for further development and testing of Pregnant Faster while searching for a more effective way to recruit the intended target group for the planned larger cohort. Another social media campaign was conducted, and all women who were previously excluded based on their neighborhood's median income were invited to participate.

Between May 2022 and July 2022, 71 (21%) more women registered for the study, adding up to the total of 337 registrations. Registration was closed after no new women registered for 2 weeks. From May onward, 31 participants were included, adding up to a total of 47 inclusions, of which 39 (83%) completed the intervention and 8 (17%) dropped out. Of these 8, 3 (37.5%) women who provided informed consent did not respond to instructions on how to install the app or attempts to reach them; 2 (25%) women were unable to install the app: in one case, an iPhone with Belgian settings preventing download from the Dutch Apple Store, and in the other case, the woman who had an Android device was scared by the warning prompted by download of an app outside of Google Play Store. Of the 3 (37.5%) remaining dropouts, 2 (25%) stopped trying to become pregnant and 1 (12.5%) found the questionnaires too burdensome. Table S1 in Multimedia Appendix 5 provides an overview of the eligibility, inclusion, and intervention completion percentages.

PCC Consultations

A total of 17 (43.6%) of 39 participants registered for a PCC consultation, and 16 (41%) consultations were conducted by 9 midwifery practices. One consultation was performed via telehealth by the first author because the midwife chosen by the participant had no experience in providing PCC and did not wish to implement PCC in her practice. The participant who registered for PCC but did not attend a consultation was worried about her health care insurance not covering the costs. Stating to feel overwhelmed, she declined additional support as well as a free telehealth consultation.

The most often reported reason to register for a PCC consultation was to obtain more personalized information (14/17, 82.4%), followed by being curious about what a consultation entails in practice (6/17, 35.3%). The most frequent reason not to register for a consultation was simply not being interested in doing so (6/22, 27.3%). A visual overview of participants' motivation regarding registration for PCC can be found in Figures S7 and S8 in Multimedia Appendix 6.

All participants who visited a PCC consultation agreed that registering through Pregnant Faster is easy (2/16, 12.5%, agree; 14/16, 87.5%, strongly agree) and were glad they had done so (3/16, 18.7%, agree; 13/16, 81.3%, strongly agree).

Coins and Rewards

During the study, participants could earn a maximum amount of 468 coins. Together, they earned a total of 11.791 coins (mean 284, SD 109 per participant; median 276, IQR 221-358; range 79-443). In total, 344 rewards were ordered during the study period (mean 8, SD 6 per participant; median 7, IQR 3-12; range 0-22). One participant did not wish to order rewards because she was happy with "just the app." She stated that she did not feel it was morally objectionable to be rewarded but just that she was not interested in receiving rewards.

The most popular reward was a €10 (US $10) book voucher, with 19 (48.7%) of 39 participants ordering the voucher at least once. The second-most popular reward was fruit, with 17 (43.6%) participants ordering fruit at least once and a total of 87 orders (87/344, 25.3%). Bananas were the most popular fruit, amounting to 32 (36.8%) of 87 fruit orders. The third-most popular reward was a set of 2 home pregnancy tests, with 16 (41%) participants ordering this reward at least once. Most participants ordered a reward more than once, displaying their personal preferences and satisfaction regarding their previous order. Table S2 in Multimedia Appendix 5 displays all rewards and their order frequency and percentage.

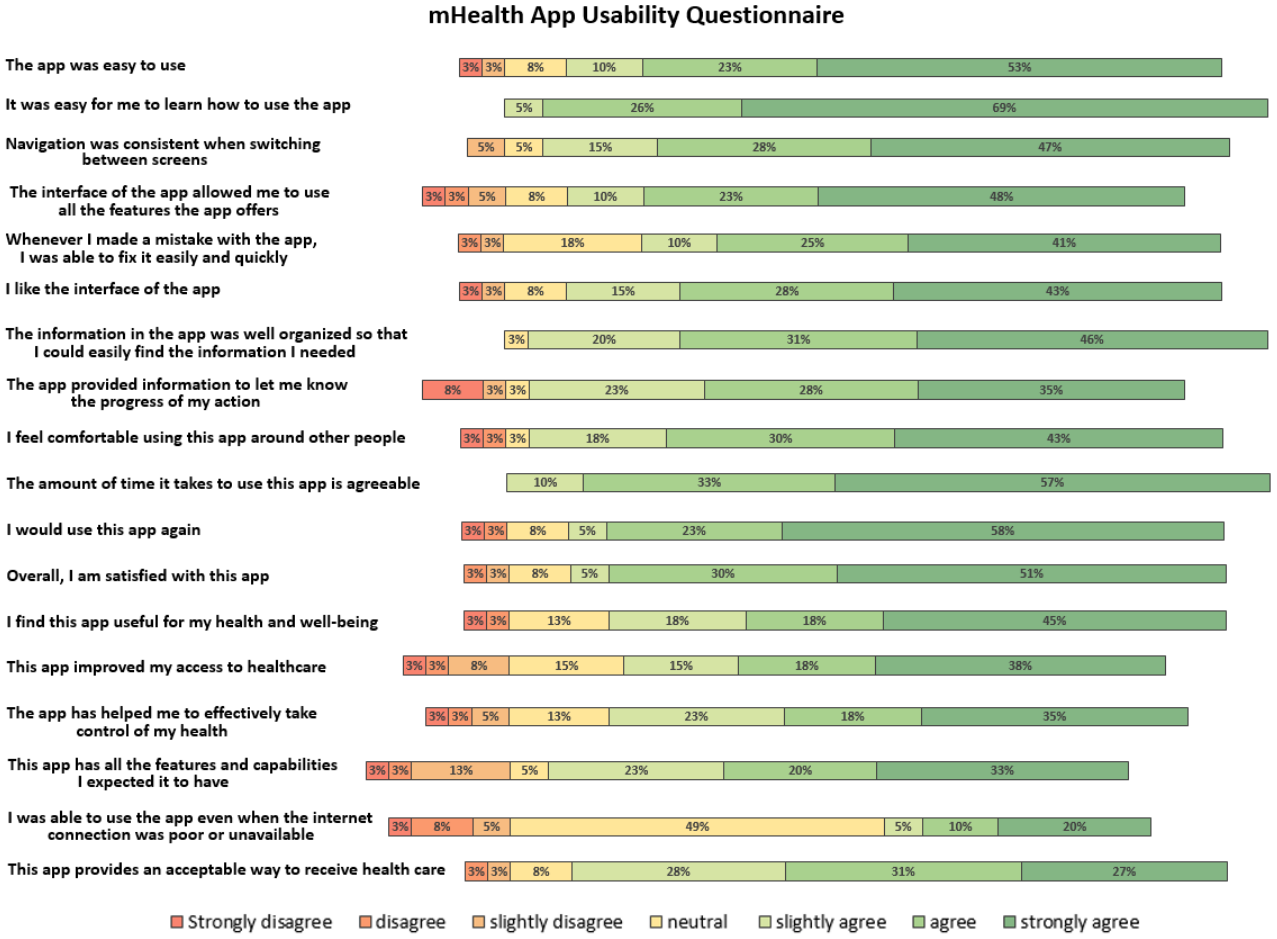
Feasibility From Users' Point of View

mHealth App Usability Questionnaire

Figure 5 displays the results of the MAUQ. The participants deemed Pregnant Faster's usability satisfying, with the majority of participants agreeing with 16 (88.9%) of 18 statements. With regard to the remaining 2 statements, all participants (N=39, 100%) agreed that the amount of time the app takes is agreeable, and 19 (48.7%) were neutral about being able to use the app with a poor internet connection, indicating they may not have experienced connectivity problems.

XSL•FO
RenderX

Figure 5. Results of the MAUQ. MAUQ: mHealth App Usability Questionnaire; mHealth: mobile health.



Experience Questionnaire

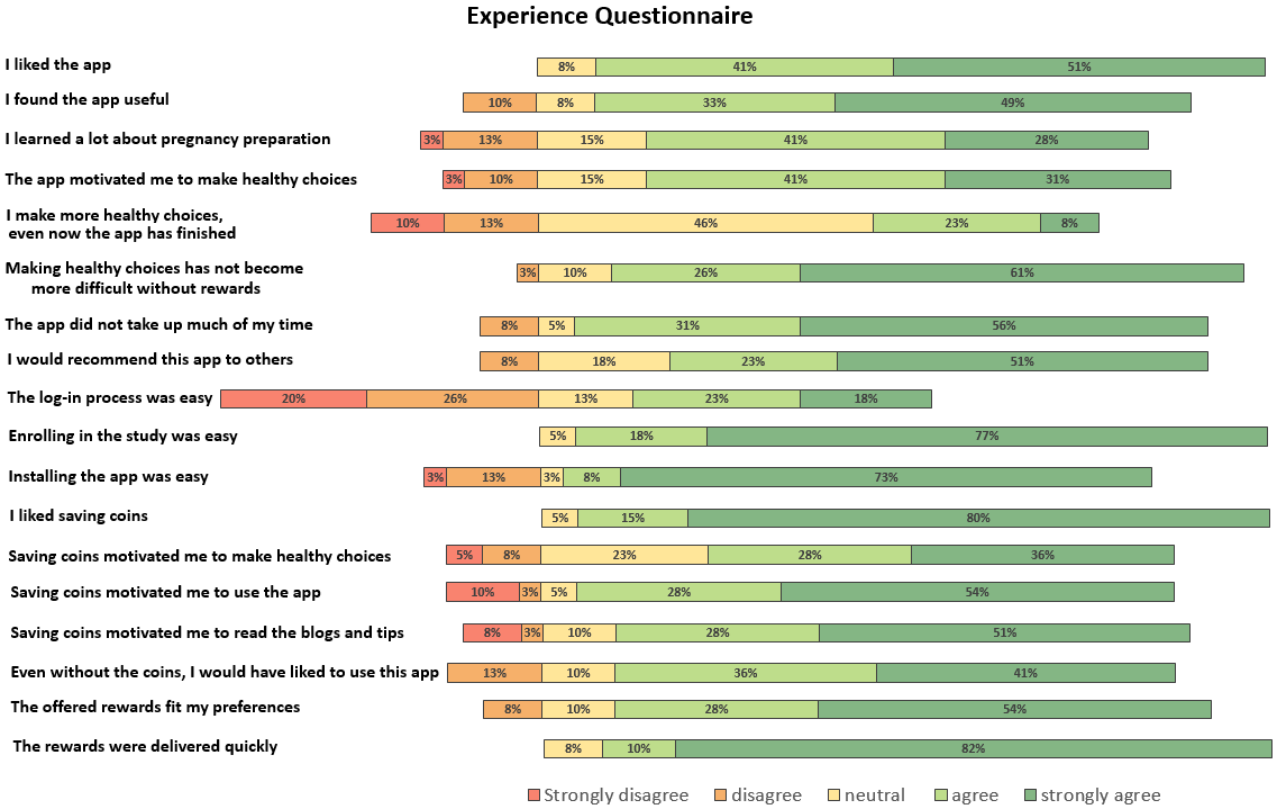
Figure 6 displays the results of the ExQ for which the 5-point Likert scale was used. The majority of participants agreed with 16 (88.9%) of 18 statements, conveying high user satisfaction. Regarding the log-in process, 16 (41%) of 39 participants agreed and 18 (46.2%) disagreed that it is easy, with 5 (12.8%) being neutral. In addition, 12 (30.8%) participants agreed and 9 (23.1%) disagreed with the statement regarding participants making more healthy choices after finishing the intervention. The remaining 18 (46.2%) participants were neutral.

The majority of participants stated that they used the app daily (n=18, 46.2%) or every other day (n=18, 46.2%). Most participants (n=31, 79.5%) reported to have logged in less often than they would have wanted to, the foremost reason being the requirement to log in with an email and password each time (n=14, 35.9%). Overall, participants rated Pregnant Faster 7.7 out of 10, with 10 being the best rating (mean 7.7, SD 1.0; median 8, IQR 7-8; range 5-9). Multimedia Appendix 6 contains Figures S9-S11, which provide additional results for the ExQ multiple-choice questions.

In the ExQ, participants were given the option to provide additional comments. It was notable that 8 (20.5%) participants commented that they would have liked push notifications to remind them of filling in the daily questionnaire and reading new blogs and tips. One participant recommended personalized notification settings so they would best fit her wishes regarding the subject, timing, and frequency. Furthermore, 3 (7.7%) participants commented that they would like the app to focus on their partners as well, hoping to actively involve them more in preparing for pregnancy.

XSL•FO
RenderX

Figure 6. Results of the ExQ. ExQ: Experience Questionnaire.



Feasibility From Researchers' Point of View

During the study, 11 issues were noted that should be considered before attempting to establish Pregnant Faster's (cost-)effectiveness in a larger cohort study:

- A time gap between the start of recruitment and intervention led to a significant loss of eligible participants and should be prevented.
- Limiting the study population to a local area greatly impedes inclusion rates and causes disappointment in otherwise eligible participants, which may harm the intervention's reputation.
- Using a combination of the neighborhood median income and a low-to-intermediate educational level as a proxy of low SES is not a suitable method to recruit large numbers of women with a VHS.
- Manual selection and inclusion require a significant amount of labor for which multiple researchers have to be available. The same goes for approaching individual midwifery practices for collaboration.
- Use of a classic, relatively complicated information folder and informed consent form can be overwhelming and does not lead to proper understanding of the study nor true consent.
- For participants with Android devices, installation of the app is complicated by not providing the app through the Google Play Store. Additional support is often needed.
- Fruit sent through the postal service often arrive bruised, requiring frequent checks for whether the reward is delivered in good condition, offering refunds when this is not the case. Failed deliveries result in the return of rotten fruit to the researchers. Since fruit is a popular, healthy reward, a suitable alternative should be considered, such as a voucher.
- Sending rewards daily is laborious. During this pilot, we experimented with a frequency of twice per week, clearly communicating this to participants. Afterward, no dissatisfaction regarding delivery time was noticed.
- Confirmation of PCC consultations requires the researchers to contact midwifery practices, causing a delay in coin allocation, possibly negatively impacting user satisfaction and effectiveness of the reward. Relying on participants self-reporting their visit in the app, combined with automatic coin allocation, should be considered.
- Manually keeping track of booked and confirmed consultations, in addition to manual coin allocation, requires a significant amount of time. For this reason as well, self-reporting should be considered.
- Asking participants to fill in 2 separate questionnaires causes confusion and diminishes the likeliness of completing the evaluation. It is advisable to evaluate user experiences in a succinct manner.

Discussion

Principal Findings

In this pilot study, we aimed to determine the feasibility of the app-based nudge Pregnant Faster, which is designed to fit the needs of women with a low SES and a high likelihood of having a VHS, who have a higher risk of adverse pregnancy outcomes. The aim of Pregnant Faster is to encourage these women by nudging them to adequately prepare for pregnancy through

XSL•FO
RenderX

education by making healthy lifestyle choices and engaging in PCC, which will help improve their pregnancy outcomes.

Pregnant Faster has shown to be feasible from the users' point of view, showing high user satisfaction with a rating of 7.7/10 and PCC uptake by 16 (41%) of 39 participants. Notably, 27 (69.2%) participants stated to have learned a lot about pregnancy preparation and 28 (71.7%) felt motivated by the app to make healthy lifestyle choices. After the intervention ended, 12 (30.8%) participants stated that they more often make healthy choices than prior to using Pregnant Faster.

With regard to the 55 (53.9%) of 102 eligible participants who did not respond when inclusion commenced after the 2-month delay, we suspect that a loss of interest and perhaps of trust in the intervention played a role. The amount of lost eligible women suggests that time is a limiting factor, impacting women's willingness to participate in the intervention. This emphasizes the necessity of quickly responding to their willingness to participate and acceptance of offered care.

Feasibility from the researchers' point of view was satisfactory as well but only with regard to practical conduction, as adjustments to the inclusion criteria were made to up the number of inclusions. Dropping the criterion of living in a deprived neighborhood likely impacted the chance of including women who actually have a VHS. The feasibility of Pregnant Faster from the researchers' point of view can be improved by developing a new method of finding the target group, making the app available via both Apple App Store and Google Play Store, and automating (parts of) the inclusion process and coin allocation, which will limit the number of administrative tasks.

Strengths and Limitations

To the best of our knowledge, Pregnant Faster is the first mHealth intervention that aims to encourage adequate pregnancy preparation and increase the uptake of PCC, promoting blended lifestyle care, by nudging participants with a loyalty program consisting of earning and spending coins. During study conduction and after evaluation, important knowledge was gained concerning the strengths and limitations of this intervention and how best to proceed with a larger cohort study.

Despite our earlier experiences regarding the recruitment of women who likely have a VHS, we did not manage to conduct this study adhering to the original inclusion criteria [28]. It is possible, therefore, that the user feasibility would have been different had the full study population met the intended criteria.

Pregnant Faster has been designed through iterative cocreation, actively involving the target population in its development [28]. Even though adjustments were made, we consider this pilot study to be another step in the iterative cocreation process, as the results will be used for further development of the app and nearly half (19/39, 48.7%) of the study population met the criteria of living in a deprived neighborhood.

Further Development and Future Research

The insights gained through this study have prompted us to re-evaluate which characteristics to use as a proxy for a low SES and the associated VHS. To improve recruitment of the target group, we have hosted meetings with health care professionals specializing in health-related vulnerability and adverse pregnancy outcomes to gain more insight and develop new inclusion criteria for a larger cohort study. At this moment, we are researching (combinations of) different inclusion criteria based on self-reported vulnerability markers, such as high stress; financial insecurity; addiction to alcohol, drugs, and tobacco; and lack of social support, which are also associated with unfavorable health outcomes [36]. Through developing these new criteria, we aim to be more inclusive and provide support to all women with a certain degree of health-related vulnerability, instead of limiting support to those with a high likelihood of having a VHS based on the educational level and neighborhood deprivation.

We aim to continue promoting Pregnant Faster on social media platforms, such as Instagram and YouTube. These platforms have been known to use algorithms that successfully reach target audiences and prove to be effective tools with regard to providing people with support and education [37,38]. Using these platforms, therefore, will not only support recruitment and benefit the target population but also allow Pregnant Faster to contribute to pregnancy-related health in the general population.

The knowledge gained through this pilot study has inspired us to research different methods of information transfer to ensure the app fits the needs of the target group and improve Pregnant Faster's accessibility for those who experience limited literacy [39]. For the planned cohort study, for example, we have created an audio version and infographic of the patient information folder and informed consent form. Furthermore, we are currently creating additional content for the app, again focusing on multiple methods of information transfer, such as podcasts, videos, and infographics.

On a technical level, the inclusion process and content management system will be adjusted to reduce manual tasks and promote feasibility. Furthermore, we plan to change the log-in procedure to a pin code or fingerprint and enhance the app with daily notifications.

Regarding focusing more on participants' partners, we have chosen to not adhere to this suggestion at the current time, as the tips and blogs already contain information for partners and we are still in the process of establishing (cost-)effectiveness and further developing Pregnant Faster. In research concerning reproductive health, it is known that partners may sometimes take on a more passive role [40], which places the burden of preparing for pregnancy largely on the person who will carry the baby. For future development, therefore, we will consider the possibility of adding personalized settings to allow users to fill in characteristics that will adjust the app's content accordingly, such as relationship status, gender and sexual orientation, and, if applicable, gestational age and the use of donor semen.

In the future, we wish to investigate the possibility of offering Pregnant Faster to all who wish to become pregnant, possibly with rewards if cost-effectiveness is established.

XSL•FO
RenderX

Conclusion

With this pilot study, we have demonstrated that the app-based nudge Pregnant Faster provides a feasible way to stimulate the uptake of PCC and boost participants' motivation to adequately prepare for pregnancy. We will use the knowledge we have gained through this pilot study to create an updated version of the app, which will be named Pregnant Faster 2. Our next step consists of determining the (cost-)effectiveness of Pregnant Faster 2, for which we will conduct a cohort study of 400 women with a VHS based on newly devised inclusion criteria.

Acknowledgments

This research was funded by ZonMw (ZorgOnderzoek Nederland en het gebied van Medische wetenschappen, health care research in the Netherlands and the field of medical sciences; grant number 543003103). We would like to thank all women who registered for this study and especially those who participated in it. Additionally, we would like to thank all midwifery practices who performed or were willing to perform PCC consultations for this study; Anneroos van der Zande, who helped continue the inclusion process and app development during the first author's maternity leave; and Maaike Stift, owner of EAT.PEASY!, who sponsored this intervention by providing vegan meals as a reward. Lastly, we thank Djanifa da Conceicao, midwife, influencer, and owner of Verlosmoeder, who contributed greatly to our inclusion process. The table-of-content image was provided by Keira Burton via Pexels.

Data Availability

The data sets used in this study are available from the corresponding author upon reasonable request.

Authors' Contributions

RST conceived the study and gained funding. RST, SS, and BB initiated the study design. SS wrote the original draft, and BB, HIM, MS, and RST reviewed and edited the original draft. BB, RST, HIM, and MS supervised the study. All authors have contributed to refinement of the study protocol and approved the final manuscript.

Conflicts of Interest

RST is initiator and developer of the mobile health (mHealth) app Smarter Pregnancy and the ZonMw project "A Loyalty Program to Motivate Vulnerable Women to Engage in Preconception Care: From Voucher to Tablet" (reference number 543003103). Other authors declare that they have no competing interests. The funder had no role in the design of the study, the writing of the manuscript, or the decision to publish the protocol.

Multimedia Appendix 1
Example of a Pregnant Faster blog.
[DOCX File , 68 KB - humanfactors_v11i1e53614_app1.docx]

Multimedia Appendix 2
The mHealth App Usability Questionnaire, adjusted for Pregnant Faster. mHealth: mobile health.
[DOCX File , 26 KB - humanfactors_v11i1e53614_app2.docx]

Multimedia Appendix 3
The Experience interview topic list.
[DOCX File , 26 KB - humanfactors_v11i1e53614_app3.docx]

Multimedia Appendix 4
The Experience Questionnaire.
[DOCX File , 27 KB - humanfactors_v11i1e53614_app4.docx]

Multimedia Appendix 5
Additional tables.
[DOCX File , 30 KB - humanfactors_v11i1e53614_app5.docx]

Multimedia Appendix 6
Additional figures.
[DOCX File , 65 KB - humanfactors_v11i1e53614_app6.docx]

XSL•FO
RenderX

References

1. Timmermans S, Bonsel GJ, Steegers-Theunissen RPM, Mackenbach JP, Steyerberg EW, Raat H, et al. Individual accumulation of heterogeneous risks explains perinatal inequalities within deprived neighbourhoods. Eur J Epidemiol 2011 Feb;26(2):165-180 [FREE Full text] [doi: 10.1007/s10654-010-9542-5] [Medline: 21203801]
2. Hulshof KFAM, Brussaard JH, Kruizinga AG, Telman J, Löwik MRH. Socio-economic status, dietary intake and 10 y trends: the Dutch National Food Consumption Survey. Eur J Clin Nutr 2003 Jan 27;57(1):128-137. [doi: 10.1038/sj.ejcn.1601503] [Medline: 12548307]
3. Beenackers MA, Kamphuis CB, Giskes K, Brug J, Kunst AE, Burdorf A, et al. Socioeconomic inequalities in occupational, leisure-time, and transport related physical activity among European adults: a systematic review. Int J Behav Nutr Phys Act 2012 Sep 19;9(1):116 [FREE Full text] [doi: 10.1186/1479-5868-9-116] [Medline: 22992350]
4. Steptoe A, Feldman PJ. Neighborhood problems as sources of chronic stress: development of a measure of neighborhood problems, and associations with socioeconomic status and health. Ann Behav Med 2001;23(3):177-185. [doi: 10.1207/S15324796ABM2303_5] [Medline: 11495218]
5. Poeran J, Maas AFG, Birnie E, Denktas S, Steegers EAP, Bonsel GJ. Social deprivation and adverse perinatal outcomes among Western and non-Western pregnant women in a Dutch urban population. Soc Sci Med 2013 Apr;83:42-49. [doi: 10.1016/j.socscimed.2013.02.008] [Medline: 23465203]
6. Vos AA, Posthumus AG, Bonsel GJ, Steegers EAP, Denktaş S. Deprived neighborhoods and adverse perinatal outcome: a systematic review and meta-analysis. Acta Obstet Gynecol Scand 2014 Aug;93(8):727-740 [FREE Full text] [doi: 10.1111/aogs.12430] [Medline: 24834960]
7. Waelput AJM, Sijpkens MK, Lagendijk J, van Minde MRC, Raat H, Ernst-Smelt HE, et al. Geographical differences in perinatal health and child welfare in the Netherlands: rationale for the healthy pregnancy 4 all-2 program. BMC Pregnancy Childbirth 2017 Aug 01;17(1):254 [FREE Full text] [doi: 10.1186/s12884-017-1425-2] [Medline: 28764640]
8. Agyemang C, Vrijkotte TGM, Droomers M, van der Wal MF, Bonsel GJ, Stronks K. The effect of neighbourhood income and deprivation on pregnancy outcomes in Amsterdam, the Netherlands. J Epidemiol Community Health 2009 Sep;63(9):755-760. [doi: 10.1136/jech.2008.080408] [Medline: 19679715]
9. Metcalfe A, Lail P, Ghali WA, Sauve RS. The association between neighbourhoods and adverse birth outcomes: a systematic review and meta-analysis of multi-level studies. Paediatr Perinat Epidemiol 2011 May;25(3):236-245. [doi: 10.1111/j.1365-3016.2011.01192.x] [Medline: 21470263]
10. Steegers-Theunissen RPM, Twigt J, Pestinger V, Sinclair KD. The periconceptional period, reproduction and long-term health of offspring: the importance of one-carbon metabolism. Hum Reprod Update 2013;19(6):640-655 [FREE Full text] [doi: 10.1093/humupd/dmt041] [Medline: 23959022]
11. Godfrey KM, Lillycrop KA, Burdge GC, Gluckman PD, Hanson MA. Epigenetic mechanisms and the mismatch concept of the developmental origins of health and disease. Pediatr Res 2007 May;61(5 Pt 2):5R-10R. [doi: 10.1203/pdr.0b013e318045bedb] [Medline: 17413851]
12. Hanson M, Godfrey KM, Lillycrop KA, Burdge GC, Gluckman PD. Developmental plasticity and developmental origins of non-communicable disease: theoretical considerations and epigenetic mechanisms. Prog Biophys Mol Biol 2011 Jul;106(1):272-280 [FREE Full text] [doi: 10.1016/j.pbiomolbio.2010.12.008] [Medline: 21219925]
13. van Zundert SKM, van Rossem L, Willemsen SP, van der Meer L, Ernst-Smelt HE, Steegers-Theunissen RPM. Periconceptional maternal social, lifestyle and medical risk factors impair embryonic growth: the Rotterdam Periconceptional Cohort. Reprod Biomed Online 2022 Jun;44(6):1123-1133 [FREE Full text] [doi: 10.1016/j.rbmo.2022.02.011] [Medline: 35361546]
14. van Uitert EM, Exalto N, Burton GJ, Willemsen SP, Koning AHJ, Eilers PHC, et al. Human embryonic growth trajectories and associations with fetal growth and birthweight. Hum Reprod 2013 Jul;28(7):1753-1761. [doi: 10.1093/humrep/det115] [Medline: 23569080]
15. de Mendonça ELSS, de Lima Macêna M, Bueno N, de Oliveira ACM, Mello C. Premature birth, low birth weight, small for gestational age and chronic non-communicable diseases in adult life: a systematic review with meta-analysis. Early Hum Dev 2020 Oct;149:105154 [FREE Full text] [doi: 10.1016/j.earlhumdev.2020.105154] [Medline: 32799034]
16. Lastrucci V, Lorini C, Caini S, Florence Health Literacy Research Group, Bonaccorsi G. Health literacy as a mediator of the relationship between socioeconomic status and health: a cross-sectional study in a population-based sample in Florence. PLoS One 2019 Dec 23;14(12):e0227007 [FREE Full text] [doi: 10.1371/journal.pone.0227007] [Medline: 31869381]
17. M'hamdi HI, Sijpkens MK, de Beaufort I, Rosman AN, Steegers EA. Perceptions of pregnancy preparation in women with a low to intermediate educational attainment: a qualitative study. Midwifery 2018 Apr;59:62-67. [doi: 10.1016/j.midw.2018.01.004] [Medline: 29396381]
18. Wildschut HIJ, van Vliet-Lachotzki EH, Boon BM, Lie Fong S, Landkroon AP, Steegers EAP. [Preconception care: an essential part of the care for mother and child]. Ned Tijdschr Geneeskd 2006 Jun 17;150(24):1326-1330. [Medline: 16808362]
19. de Weerd S, Steegers EAP, Heinen MM, van den Eertwegh S, Vehof RMEJ, Steegers-Theunissen RPM. Preconception nutritional intake and lifestyle factors: first results of an explorative study. Eur J Obstet Gynecol Reprod Biol 2003 Dec 10;111(2):167-172. [doi: 10.1016/s0301-2115(03)00290-2] [Medline: 14597246]

XSL•FO
RenderX

20. Van Dijk MR, Huijgen NA, Willemsen SP, Laven JS, Steegers EA, Steegers-Theunissen RP. Impact of an mHealth platform for pregnancy on nutrition and lifestyle of the reproductive population: a survey. JMIR Mhealth Uhealth 2016 May 27;4(2):e53 [FREE Full text] [doi: 10.2196/mhealth.5197] [Medline: 27234926]
21. Poels M, Koster MPH, Boeije HR, Franx A, van Stel HF. Why do women not use preconception care? A systematic review on barriers and facilitators. Obstet Gynecol Surv 2016 Oct;71(10):603-612. [doi: 10.1097/OGX.0000000000000360] [Medline: 27770130]
22. van Dijk MR, Koster MPH, Oostingh EC, Willemsen SP, Steegers EAP, Steegers-Theunissen RPM. A mobile app lifestyle intervention to improve healthy nutrition in women before and during early pregnancy: single-center randomized controlled trial. J Med Internet Res 2020 May 15;22(5):e15773 [FREE Full text] [doi: 10.2196/15773] [Medline: 32412417]
23. Gootjes DV, van Dijk MR, Koster MP, Willemsen SP, Steegers EA, Steegers-Theunissen RP. Neighborhood deprivation and the effectiveness of mobile health coaching to improve periconceptional nutrition and lifestyle in women: survey in a large urban municipality in the Netherlands. JMIR Mhealth Uhealth 2019 Apr 11;7(4):e11664 [FREE Full text] [doi: 10.2196/11664] [Medline: 30973345]
24. Smith SM, van der Kleij RMJJ, Bais B, Schermer MHN, Ismaili M'hamdi H, Steegers-Theunissen RPM. Preferences of women with a vulnerable health status towards nudging for adequate pregnancy preparation as investment in health of future generations: a qualitative study. BMC Pregnancy Childbirth 2022 Jul 13;22(1):559 [FREE Full text] [doi: 10.1186/s12884-022-04887-6] [Medline: 35831786]
25. Thaler R, Sunstein C. Nudge : improving decisions about health, wealth, and happiness. New Haven, Conn. New Haven, CN: Yale University Press; 2008.
26. Murayama H, Takagi Y, Tsuda H, Kato Y. Applying nudge to public health policy: practical examples and tips for designing nudge interventions. Int J Environ Res Public Health 2023 Feb 23;20(5):3962 [FREE Full text] [doi: 10.3390/ijerph20053962] [Medline: 36900972]
27. van der Windt M, van der Kleij RM, Snoek KM, Willemsen SP, Dykgraaf RHM, Laven JSE, et al. Impact of a blended periconception lifestyle care approach on lifestyle behaviors: before-and-after study. J Med Internet Res 2020 Sep 30;22(9):e19378 [FREE Full text] [doi: 10.2196/19378] [Medline: 32996885]
28. Smith SM, Bais B, Ismaili M'hamdi H, Schermer MHN, Steegers-Theunissen RPM. Stimulating the uptake of preconception care by women with a vulnerable health status through mHealth app-based nudging (Pregnant Faster): cocreation design and protocol for a cohort study. JMIR Res Protoc 2023 Aug 09;12:e45293 [FREE Full text] [doi: 10.2196/45293] [Medline: 37556197]
29. Singh K, Rishiwal V, Kumar P. Classification of data to enhance data security in cloud computing. 2018 Presented at: IoT-SIU 2018: 3rd International Conference On Internet of Things: Smart Innovation and Usages; February 23-24, 2018; Bhimtal, India p. 23-24. [doi: 10.1109/iot-siu.2018.8519934]
30. Would you like to become pregnant? Sneller Zwanger. URL: https://www.snellerzwangerworden.nl/nl/sneller-zwanger [accessed 2024-04-11]
31. VerlosMoeder online platform. VerlosMoeder. URL: https://www.verlosmoeder.nl/ [accessed 2024-04-12]
32. VerlosMoeder. YouTube. URL: https://www.youtube.com/@VerlosMoeder [accessed 2024-04-12]
33. Djanifa da Conceicao. Instagram. URL: https://www.instagram.com/Djanifa_VerlosMoeder/ [accessed 2024-04-12]
34. Smarter Pregnancy: for couples who are or wish to get pregnant. Erasmus University Medical Center. 2019. URL: https://www.smarterpregnancy.co.uk/ [accessed 2024-04-11]
35. Zhou L, Bao J, Setiawan IMA, Saptono A, Parmanto B. The mHealth App Usability Questionnaire (MAUQ): development and validation study. JMIR Mhealth Uhealth 2019 Apr 11;7(4):e11500 [FREE Full text] [doi: 10.2196/11500] [Medline: 30973342]
36. Aday LA. Health status of vulnerable populations. Annu Rev Public Health 1994 May;15:487-509. [doi: 10.1146/annurev.pu.15.050194.002415] [Medline: 8054096]
37. Carpenter J, Morrison S, Craft M, Lee M. How and why are educators using Instagram? Teach Teach Educ 2020 Nov;96:103149 [FREE Full text] [doi: 10.1016/j.tate.2020.103149] [Medline: 32834464]
38. Shoufan A, Mohamed F. YouTube and education: a scoping review. IEEE Access 2022;10:125576-125599. [doi: 10.1109/access.2022.3225419]
39. Seligman HK, Wallace AS, DeWalt DA, Schillinger D, Arnold CL, Shilliday BB, et al. Facilitating behavior change with low-literacy patient education materials. Am J Health Behav 2007 Jul 01;31(1):69-78. [doi: 10.5993/ajhb.31.s1.9]
40. Pachauri S. Male involvement in reproductive health care. J Indian Med Assoc 2001 Mar;99(3):138-141. [Medline: 11478756]

Abbreviations

CIA: confidentiality, integrity, and availability
ExQ: Experience Questionnaire
MAUQ: mHealth App Usability Questionnaire
mHealth: mobile health
PCC: preconception care

XSL•FO
RenderX

SES: socioeconomic status
VHS: vulnerable health status

Edited by A Choudhury; submitted 12.10.23; peer-reviewed by S Verbiest, R Marshall; comments to author 14.02.24; revised version received 22.02.24; accepted 26.02.24; published 22.04.24.

Please cite as:
Smith SM, Bais B, Ismaili M'hamdi H, Schermer MHN, Steegers-Theunissen RPM
Stimulating Preconception Care Uptake by Women With a Vulnerable Health Status Through a Mobile Health App (Pregnant Faster): Pilot Feasibility Study
JMIR Hum Factors 2024;11:e53614
URL: https://humanfactors.jmir.org/2024/1/e53614
doi:10.2196/53614
PMID:38648092

©Sharissa M Smith, Babette Bais, Hafez Ismaili M'hamdi, Maartje HN Schermer, Régine PM Steegers-Theunissen. Originally published in JMIR Human Factors (https://humanfactors.jmir.org), 22.04.2024. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on https://humanfactors.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Efficacy of the QuitSure App for Smoking Cessation in Adult Smokers: Cross-Sectional Web Survey

Gregory M Goldgof[1*], MD, PhD; Shweta Mishra[2], MA; Kriti Bajaj[2*], MSc

[1]Department of Laboratory Medicine, University of California, San Francisco, San Francisco, CA, United States

[2]QuitSure, Rapidkart Online Private Limited, Mumbai, India

[*]these authors contributed equally

Corresponding Author:
Kriti Bajaj, MSc
QuitSure
Rapidkart Online Private Limited
Bajaj Bhavan, 2nd Floor
226 Nariman Point
Mumbai, 400021
India
Phone: 91 2222023626
Email: kriti@quitsure.app

Abstract

Background: Cigarette smoking remains one of the leading causes of preventable death worldwide. A worldwide study by the World Health Organization concluded that more than 8 million people die every year from smoking, tobacco consumption, and secondhand smoke. The most effective tobacco cessation programs require personalized human intervention combined with costly pharmaceutical supplementation, making them unaffordable or inaccessible to most tobacco users. Thus, digital interventions offer a promising alternative to these traditional methods. However, the leading smartphone apps available in the market today have either not been studied in a clinical setting or are unable to match the smoking cessation success rates of their expensive offline counterparts. We would like to understand whether QuitSure, a novel smoking cessation app built by Rapidkart Online Private Limited, is able to bridge this efficacy gap and deliver affordable and effective smoking cessation at scale.

Objective: Our objective was to do an initial exploration into the engagement, efficacy, and safety of QuitSure based on the self-reported experiences of its users. Outcomes measured were program completion, the effect of program completion on smoking behavior, including self-reported cessation outcomes, and negative health events from using the app.

Methods: All QuitSure registered users who created their accounts on the QuitSure app between April 1, 2021, and February 28, 2022, were sent an anonymized web-based survey. The survey results were added to their engagement data on the app to evaluate the feasibility and efficacy of the app as a smoking cessation intervention. The data were analyzed using descriptive statistics (frequencies and percentages) and the χ^2 test of independence.

Results: In total, 1299 users who had completed the QuitSure program submitted the survey and satisfied the inclusion criteria of the study. Of these, 1286 participants had completed the program more than 30 days before filling out the survey, and 1040 (80.1%, 95% CI 79.1%-82.6%) of them had maintained prolonged abstinence for at least 30 days after program completion. A majority of participants (770/891, 86.4%) who were still maintaining abstinence at the time of submitting the survey did not experience any severe nicotine withdrawal symptoms, while 41.9% (373/891) experienced no mild withdrawal symptoms either. Smoking quantity prior to completing the program significantly affected quit rates ($P<.001$), with heavy smokers (>20 cigarettes per day) having a lower 30-day prolonged abstinence rate (relative risk=0.91; 95% CI 90.0%-96.2%) compared to lighter smokers. No additional adverse events outside of known nicotine withdrawal symptoms were reported.

Conclusions: The nature of web-based surveys and cohort selection allows for extensive unknown biases. However, the efficacy rates of survey respondents who completed the program were high and provide a case for further investigation in the form of randomized controlled trials on the QuitSure tobacco cessation program.

(JMIR Hum Factors 2024;11:e49519) doi:10.2196/49519

XSL•FO
RenderX

KEYWORDS

smoking; quit smoking; smoking cessation; smoking app; QuitSure; smoke free; quit vaping; vaping; smoker; smoke; cross-sectional study; smartphone app; tobacco consumption; tobacco; survey; nicotine; nicotine withdrawal; mobile phone

Introduction

Background

Cigarette smoking remains one of the leading causes of many premature deaths worldwide [1]. According to the World Health Organization (WHO), more than 8 million people around the world die every year, either directly or indirectly (via secondhand smoke), because of tobacco consumption. Hence, the WHO [2] has identified the tobacco epidemic as one of the world's biggest public health threats. Beyond the burden of mortality lies the burden of disease as a result of tobacco consumption. For every 1 individual who dies because of smoking, at least 30 live with a serious illness caused by smoking. Smoking causes many health issues, such as cardiovascular diseases, chronic obstructive pulmonary disease, and 12 types of cancer [3]. More than 67% of smokers face debilitating, and eventually fatal, health issues at some point in their smoking lives [4]. Health risks, as well as death risks for smokers compared to nonsmokers, have worsened, due to the deadly spread of COVID-19 across the world [5]. Meanwhile, the economic costs attributable to smoking and exposure to tobacco smoke globally have been estimated to be US $1436 per smoker, which is equivalent to around 1.8% of the world's gross domestic product [3,6].

In 2015, 68% (22.7 million) of adult smokers said that they wanted to quit smoking. In 2018, 55.1% (21.5 million) of adult smokers said that they had made a quit attempt in the past year. In 2020, 62.5% of youths (middle and high school students) who currently used tobacco products wished to quit all tobacco products, and 65.4% of youths who currently used tobacco products reported that they had stopped using all tobacco products for 1 day or longer in the past year because they were trying to quit [7]. But on the other side of the coin, a report by the National Institute of Cancer, United States [8], found that in 2020, of the 53.9% of smokers who attempted to quit smoking, only 8.5% of them were successful in doing so. In fact, research has found that it takes about 30 quitting attempts for a smoker to successfully quit [9]. The WHO, in 2022, said that without cessation support, only 4% of smokers will be able to successfully quit.

Smoking Cessation

There are several smoking cessation methods available across the world, including unassisted methods, nicotine replacement therapy (NRT), prescribed medicine (bupropion or varenicline) use, behavioral counseling, quitlines, and the use of mobile apps and websites for smoking cessation [10]. Financial incentives have gained popularity as a cessation method recently [11]. NRT, like nicotine patches, gums, and nasal sprays; medications such as bupropion and varenicline; and nonpharmacological interventions [12] are the most common smoking cessation interventions. However, NRT has shown to have success rates of only 6%-8% [13], while pharmacological interventions, despite their somewhat higher success rates of 14%-20% [14], come with the risk of side effects such as skin irritation or more serious seizures and are also very expensive [15]. Combined interventions for smokers such as behavioral interventions and long-term assistance or social support are most effective when it comes to smoking cessation [15]. However, they tend to be expensive, highly variable depending on the quality of each individual provider, accessible to only small hyperlocal communities, and cannot be scaled up to achieve population-level impact.

Smartphone Apps for Smoking Cessation

In response to the COVID-19 pandemic, the rate of smoking cessation increased from 23% to 31% [16], which creates the opportunity to encourage and support smokers to quit smoking through different smoking cessation methods. Unconventional methods such as smartphone-based apps can be more useful to increase the odds of quitting success over conventional methods because smartphone use is highly prevalent, is available 24-7, is cost-effective, requires zero-minimal human intervention, and can provide instant and constant support. Seo et al [17] found a total of 603 apps designed for smoking cessation that were available in the US, UK, Australian, and Asian markets [17]. Apps designed for smoking cessation have been downloaded 33 million times globally according to a study done by SensorTower in April 2020 (Nelson, SensorTower.com, personal communication, April 15, 2020). Users who have high engagement with smoking cessation apps have been found to be more likely to be successful in quitting [18,19].

However, literature reviews suggest that only a few apps follow the guidelines for treating tobacco dependence, and most apps use only very simple tools like calculators (41%), calendars (36%), trackers (18%), hypnosis apps (21%), and distractors (10%) [20,21]. According to the WHO, any primary care provider needs to follow the 5As (ask, advise, assess, assist, and arrange) to help a tobacco user [22]. One content analysis study suggested that 96% of the cessation apps addressed "assist" but less frequently addressed the other 4 As [21]. Another review demonstrated that only 11 (6.1%) of the 180 smoking cessation apps available in 2022 have any scientific support [23]. The review also discovered that very few apps offered evidence-based interventions such as mindfulness (18%) or cognitive behavioral therapy (CBT; 2.2%). Other reviews indicated that 88.46% of smoking cessation apps have not been updated by the developers in over a year, and 33.67% of apps have low acceptance by the market with <10,000 downloads [17,21]. Thus, the development of additional smartphone apps that have good user acceptance and are using empirically supported behavior change techniques to deliver smoking cessation interventions appears to be warranted.

The QuitSure Smoking Cessation Program

The QuitSure program (Rapidkart Online Private Limited) has been identified as one such program, which incorporates behavior change techniques like positive psychology, CBT, and mindfulness that have been shown to be effective in smoking

XSL•FO
RenderX

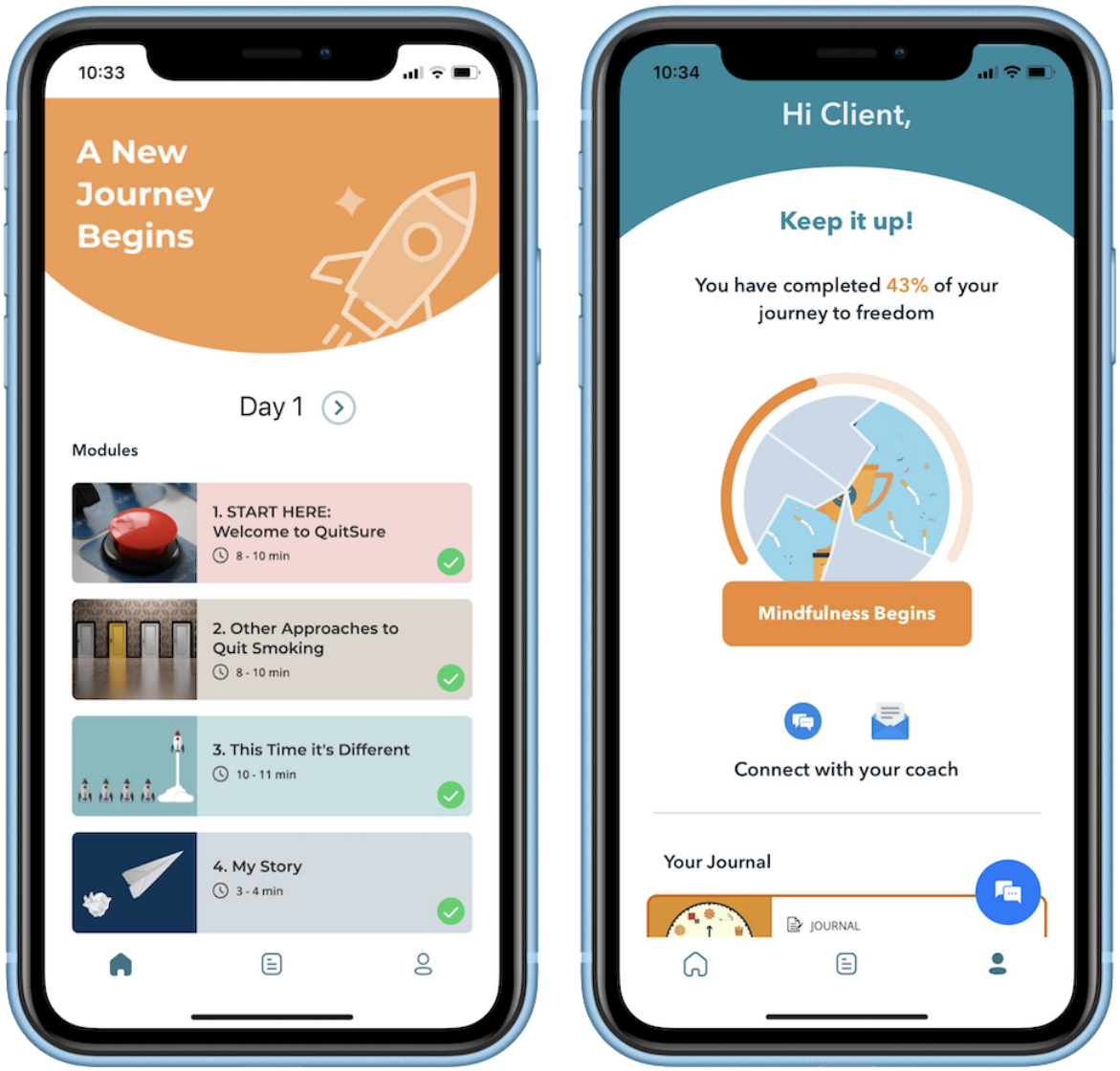
cessation interventions [22,24], and is customized to the smoking habits and psychological needs of the user. It does not include or recommend the use of any pharmaceutical interventions, like oral supplements, medications, or NRTs.

The program follows the 5As recommended by the WHO [22]: it first asks the users about their smoking history such as quantity, patterns, past quitting attempts, and relapse reasons. This also helps the user internalize their smoking behavior. Then, it assesses if they are ready to quit and asks them about their inhibitions to quit. Unless it is a medical condition, it then advises in a clear, strong manner to quit smoking alongside a summary of the program and how it actually works. After understanding the program, willing users are then progressed to the next level, where the app provides the main content on psychology, CBT, and mindfulness [24].

CBT is incorporated by helping users question their beliefs around the positive aspects of smoking and remove them. The app does not demand any lifestyle modification. It simply helps the user accept their smoking triggers and change their response to these triggers. Under mindfulness, the app provides video exercises to teach users how to smoke mindfully by focusing on every aspect of the smoking experience. This exposes to the user the real effects of cigarettes, both while smoking and after, on their bodies and minds. All the content is delivered using empathy and without administering any guilt or blame to the user to keep the user's mind in a more calm and receptive state.

Users are required to complete the program in a very specific way with a predefined sequence of content and video exercises as shown in the screenshot of the home page in Figure 1 (left). The program requires around 6-10 hours over 6-12 days to complete. During the whole process, the app has a structured, digital journal for users to record their quitting journey and 24×7 chat-based support from trained counselors for users who have additional questions or concerns that are not addressed by the program itself as shown in Figure 1 (right). Once the user has completed the program and quit smoking, the app has postquit tools and chat support available to prevent them from relapsing. Around 12.3% (4124/33,458) of all users reach out to the counselors for support during and after the program.

Figure 1. Screenshots of the home page (left) and profile page (right) of the QuitSure app.



XSL•FO
RenderX

The program is priced at US $10 and should be affordable to most daily smokers, as the average global retail price of a pack of 20 cigarettes is approximately US $5 [25,26]. Users only have to pay on day 2 of the 6-day program. This is done, so they can understand how the program is structured and what techniques it uses before committing to it. At the end of the program, users are asked to perform a guided smoking exercise where they smoke their last cigarette and then quit cold turkey. The timestamp when they mark having performed this exercise in the app is considered as their quit date and as their program completion date. The QuitSure app has been updated an average of once every 3 weeks since its launch.

In this paper, we report the results of a retrospective cross-sectional study on users who completed the program between April 1, 2021, and February 28, 2022. Participants who started the QuitSure program but then dropped out before completing it were also surveyed to evaluate the potential areas of improvement for QuitSure. The aim of this study is to evaluate the effectiveness of the QuitSure smoking cessation program to enable smoking cessation among daily smokers. We will also examine the program's usability, feasibility, and acceptance by the market.

Methods

Design

This was a retrospective cross-sectional study to understand, through a web-based survey, smoking cessation outcomes of users who completed the QuitSure program via the QuitSure smartphone app. The survey was conducted in April 2022.

Recruitment

Users who downloaded and registered for the QuitSure quit smoking app between dates April 1, 2021, and February 28, 2022, and satisfied all the following inclusion criteria were sent the web surveys.

The study included (1) users who were daily smokers as defined by the US Centers for Disease Control and Prevention—smoked at least 1 cigarette per day before they did the program and had smoked at least 100 cigarettes before doing the program [6]; (2) adults aged 18 years and older (self-reported); (3) users who were, at minimum, proficient in the English language (since the program content is solely available in English); (4) users who had completed the entire QuitSure program; (5) and users who had a valid email address.

All registered users of the QuitSure app give consent to be contacted for the purpose of research studies at the time of registration.

Web-Based Survey

The surveys were created on the web-based Jotform tool built by Jotform Inc. The forms were set up with Jotform's Health Insurance Portability and Accountability Act (HIPAA) compliant mode, and no personally identifiable data were collected, thus protecting the confidentiality and privacy of the participants. The survey details have been reported in accordance with the Checklist for Reporting Results of Internet E-Surveys (CHERRIES) guidelines [27].

The survey was sent via an automated email from the QuitSure server to all users who fit the inclusion criteria. The email included the relevant details for participation in the study, including the length of the survey, the aim of the study, benefits for the participant, as well as the link to the survey itself. They were encouraged to contact the investigators if they had any questions or concerns. Participation in the survey was voluntary. Participants were granted entry into a lucky draw for a US $50 Amazon voucher. This amount was chosen as being 5 times the price of the QuitSure program, and the email specified that winning the US $50 voucher did not depend on whether their experience with the QuitSure program was positive or negative.

To prevent duplication of data, each email included a hidden, unique, non–personally identifiable ID number that appeared in the survey results. For duplicate submissions, the more recent entry was kept. All questions seen by the participant, depending on the conditions, were mandatory. As a result, incomplete surveys were not recorded or used for analysis. No statistical correction methods, such as weighting of items or propensity scores, were applied to adjust for sample nonrepresentativeness.

The first part of the survey confirmed the participants' demographic data and satisfaction of the inclusion criteria. It also explicitly requested informed consent for participation in the study. Participants were shown the remaining sections in the survey only if they consented to participate in the study and met the inclusion criteria defined earlier. The second and third parts of each survey are described in the individual sections below. Prior to launching the survey, we conducted usability and technical functionality testing to ensure that participants could easily navigate and complete the electronic questionnaire. In both surveys, the questions were primarily multiple choice, with the rest requiring integer number entries. All multiple choice questions included the option for free-form "other" entries as well as the option for "none" or "choose not to share" as relevant for that question.

Survey for Program Completers

This survey, referred to as S-Completers, was sent to all users who satisfied the inclusion criteria and had also completed the QuitSure program on the QuitSure app as defined in the introduction. The second section of this survey identified the participants' smoking history, including the outcome of their completion of the QuitSure program. The third section was conditional. Each participant saw a different set of questions depending on their smoking cessation or reduction or relapse outcome after program completion. The number of questions in each section ranged from 4 to 9 depending on the responses of the participant.

Survey for Noncompleters of the Program

This survey, referred to as the S-Incompleters survey, was sent to all users who satisfied the inclusion criteria and who had started the program, but then not completed it, to get a qualitative understanding of the feasibility of the program at scale. The S-Incompleters survey also took the participants' explicit consent for participation in the study and the same demographic data and smoking history questions. The rest of the survey was to understand their reasons for leaving the

XSL•FO
RenderX

program midway. The number of questions in each section ranged from 4 to 9 depending on the responses of the participant. The copy of both survey questionnaires is provided in Multimedia Appendices 1 and 2.

Study Variables

The data collected in the first part of both surveys included demographic information about the participants including gender, age, country of residence, and English proficiency. The second part of both surveys, regarding smoking history and behavior, asked what forms they consumed nicotine in before doing the program, how much and how long they had smoked, and whether they had previously tried to quit smoking using other methods. Those who had completed the QuitSure program were asked for the outcome of their most recent attempt at doing the QuitSure program and if they used any other quit smoking tools, programs, or medications or supplements during or after doing the QuitSure program.

The questions in the third part of the S-Completers survey were dependent on their outcome of doing the QuitSure program. Those who were able to quit 100% since completing the program were asked whether they experienced any mild or severe withdrawal symptoms and weight gain. They were also asked for their current level of cravings to smoke via a Likert scale ranging from 1=none to 5=unbearable. Those who quit for some time, but then relapsed, were asked how long they were able to stay quit and their reasons for relapse. Those who were only able to cut down their smoking level were asked for their new smoking rates and the reasons the program did not help them quit completely. Finally, those for whom the program had no impact, were asked why the program did not work for them. All participants who said they are still smoking were asked their current level of motivation to quit smoking on the Likert scale and whether they would use QuitSure again for their next quitting attempt.

The S-Incompleters survey asked participants why they did not complete the program and whether they were able to quit using a different method since. If they have quit since, then how they quit and what their level of cravings to smoke is. If not, then what their level of motivation to quit is and whether they will use QuitSure as their method of choice.

Ethical Considerations

This study was approved by the institutional review board of the University of California, San Francisco (IRB 21-35619, reference 331340). The email sent to the users described the study's aims and procedures as well as the security and confidentiality of their data. It also clearly stated that participation was voluntary, and they could decline to participate. The participants were given a consent form with all the details of the study including the purpose of the study. Participants younger than 18 years were not included in the study. The study observed data protection laws in effect at the time it was conducted. Participants were entered in a lucky draw to win a US $50 gift voucher for Amazon.

Statistical Analysis

Data were collected from 1299 participants from over 25 countries. The survey responses were available in Microsoft Excel (Microsoft Corp) format. The data analysis tools used were descriptive statistics (frequencies and percentages), pivot tables, as well as chi-square tests of independence.

Results

Survey for Program Completers

Of the 13,585 users who were sent the S-Completers survey email, 853 (6.3%) emails bounced. Of the 12,732 delivered emails, 5365 (42.1%) were opened. QuitSure or smoking was not mentioned in the sender or subject lines of the emails, only in the body of the email, to reduce bias based on user perception of the QuitSure program before opening the email. Therefore, we will use this number of email openers as the baseline number of users who were aware of this study. In total, 1906 (35.5%) of those who opened the email clicked on the survey link, and a final 1332 (24.8%) email openers consented to participate in the study and completed the survey. These values are all either equal to or greater than the expected opening (21.5%) and click (8%) rates based on global industry standards [28,29]. Of this set of submitters, 11 were excluded for completing the program in less than 7 days before submitting the survey. An additional 22 were excluded for submitting false data that were significantly different from actual app engagement. A final set of 1299 participants were included in the data analysis for the study. App engagement and preprogram smoking behavior data were not significantly different between those who filled out the survey versus those who did not. A flowchart representing the participant funnel for the S-Completers survey is available in Multimedia Appendix 3.

Table 1 shows the baseline characteristics of the participants, while Multimedia Appendix 4 shows their global distribution. The ratio of male to female participants was found to be 1 to 1.17. The most common age range of participants was found to be 25-34 years (n=431, 33.2%). In total, 97.2% (n=1262) of participants were cigarette smokers, while the remaining 2.6% (n=37) only consumed other forms of nicotine. Five countries, the United States, the United Kingdom, India, Canada, and Australia, had 71.8% (n=933) of the participants.

XSL•FO
RenderX

Table 1. Demographic details of participants.

	Male participants (n=587)	Female participants (n=699)	Others (n=13)	All (N=1299)
Age (years), mean (SD)[a]	35.8 (11.74)	41.83 (12.55)	27.11 (6.82)	39.5 (12.56)
Cigarettes smoked per day, mean (SD)	12.57 (8.02)	14.57 (8.56)	13.30 (5.51)	13.66 (8.83)
Median (IQR; smoking category)	10 (7-20; light)	14 (10-20; average)	13 (10-18; average)	12 (8-20; average)
Years smoked, mean (SD)	15.28 (11.40)	21.13 (12.87)	10.15 (9.69)	18.59 (12.62)
Smoking categories (cigarettes per day), n (%)				
Very light (<5)	75 (12.8)	47 (6.7)	0 (0)	122 (9.4)
Light (5-10)	119 (20.2)	109 (15.6)	3 (25)	231 (17.8)
Average (11-20)	233 (39.6)	301 (43.1)	8 (61.5)	542 (41.7)
Heavy (>20)	160 (27.2)	242 (34.6)	2 (15.4)	404 (31.1)

[a]To calculate the mean, the midpoints of each age group were considered (calculated as upper limit+lower limit/2, ie, 25+35/2 and so on).

Smoking behavior of participants prior to doing the program was grouped into 4 categories as very light (<5 cigarettes per day), light (5-10 cigarettes per day), average (11-20 cigarettes per day), and heavy (>20 cigarettes per day) [30]. The participants smoked an average of 13.66 (SD 8.83) cigarettes per day.

Effectiveness of the Smartphone App for Smoking Cessation

Participants were divided into 4 overlapping subsets based on the duration between completing the program and submitting the survey. The 4 durations chosen were 7 days, 30 days, 3 months, and 6 months, which are the commonly used smoking cessation milestones [31]. Table 2 shows the self-reported outcome of completing the program for each of these groups. Overall, 88% (1144/1299), 80.9% (1040/1286), 82.4% (991/1203), and 72.4% (725/1002) of participants had maintained prolonged abstinence for 7 days, 30 days, 3 months, and 6 months, respectively. In total, 35 of the 1203 (2.9%) participants were able to cut down on their smoking level in 30 days after the completion of the program, and 19 of the 1002 (1.9%) were able to sustain it for over 6 months after program completion.

Table 2. Prolonged abstinence after program completion.

	Participants with 7-day prolonged abstinence (1144/1299, 88%)	Participants with 30-day prolonged abstinence (1040/1286, 80.9%)	Participants with 3-month prolonged abstinence (991/1203, 82.4%)	Participants with 6-month prolonged abstinence (725/1002, 72.4%)
OR[a] (95% CI)	1.32 (86.8-89.7)	1.20 (79.4-82.3)	1.39 (80-84.6)	1.43 (69.8-75.8)
Participants by smoking behavior, n/N (%)				
Very light (<5)	107/122 (87.7)	95/121 (78.5)	84/101 (83.2)	57/87 (65.5)
Light (5-10)	209/231 (90.5)	196/228 (86)	183/207 (88.4)	132/173 (76.3)
Average (11-20)	491/542 (90.6)	446/535 (83.4)	405/466 (86.9)	319/425 (75.1)
Heavy (>20)	333/404 (82.4)	301/401 (75.1)	271/340 (79.7)	215/316 (68)
Participants by country, n/N (%)				
United States	351/393 (89.3)	324/388 (83.5)	278/357 (77.9)	242/315 (76.8)
United Kingdom	88/104 (84.6)	73/103 (70.9)	67/97 (69)	56/86 (65)
India	226/266 (84.6)	195/261 (74.7)	158/231 (68.4)	129/200 (64.5)
Canada	78/87 (90)	72/86 (83)	58/75 (77)	51/67 (76)
Australia	68/75 (91)	65/75 (87)	54/67 (81)	50/59 (85)

[a]OR: odds ratio.

The participants who were able to quit smoking as a result of the program and had maintained prolonged cessation at the time of filling out the survey (n=891) experienced varying degrees of withdrawal symptoms, as shown in Tables 3 and 4. Overall, 41.9% (n=373) experienced no mild withdrawal symptoms, and 86.4% (n=770) experienced no severe withdrawal symptoms. In total, 41.9% (n=373) experienced no weight gain after quitting with the QuitSure app; 39.6% (n=353) gained less than 5 kg, while 18.5% (n=165) gained more than 5 kg after quitting.

XSL•FO
RenderX

Table 3. Mild withdrawal symptoms reported by participants (n=891).

Symptom	Participants, n (%)
No mild withdrawal symptoms	373 (41.9)
Some mood issues	349 (39.2)
Mild sleep disturbances	196 (22)
Coughing or mild nausea	167 (18.7)
Mild digestive changes	145 (16.3)
Low energy or weakness	143 (16)
Mild headaches	131 (14.7)
Tingling of hands and feet	42 (4.7)
Others	38 (4.3)

Table 4. Severe withdrawal symptoms reported by participants (n=891).

Symptom	Participants, n (%)
No severe withdrawal symptoms	770 (86.4)
Increased depression or anxiety	90 (10.1)
Severe headaches or migraines	23 (2.6)
Severe insomnia	22 (2.5)
Severe dizziness or nausea or weakness	19 (2.1)
Strong chest pain	16 (1.8)
Others	10 (1.1)

Factors Contributing to Success

To assess whether quitting smoking via QuitSure is independent of demographic variables, 2 chi-square tests of independence were conducted as shown in Table 5. The χ^2 value for the impact of gender and age on the efficacy of the program was found to be χ^2_2=3.8 (*P*=.09) and χ^2_4=5.9 (*P*=.20), respectively. This indicates that smoking cessation via QuitSure was not dependent on the gender or age of the participants. Smoking behavior prior to starting the program, however, did affect the program's efficacy. The value was found to be χ^2_1=20.3 (*P*<.001), indicating a significant impact of smoking behavior on the quit rate. The 30-day prolonged abstinence of heavy smokers was significantly lower than that of those who smoked <20 cigarettes a day (relative risk=0.91; 95% CI 90.0%-96.2%). Country of residence also had a significant impact on program effectiveness with a value of χ^2_4=9.8 (*P*=.04) when comparing the 5 countries with the most participants. Residents of Australia had the highest 30-day prolonged abstinence rates (relative risk compared to all other participants=1.08, 95% CI 44.0%-83.0%), while residents of the United Kingdom had the lowest (relative risk for 30-day prolonged abstinence compared to all other participants=0.87, 95% CI 84.0%-99.2%).

Table 5. Factors affecting smoking cessation rates.

Factor	Chi-square (*df*)	*P* value
Gender	3.8 (2)	.09
Age groups	5.9 (4)	.20
Smoking behavior	20.3 (1)	<.001
Country of residence	9.8 (4)	.04

Factors Contributing to Failure and Relapse

Table 6 shows the reasons given for failure among participants for whom the program did not work at all (n=35), with fear of quitting (n=15, 42.9%) and lack of belief (n=11, 31.4%) being the most common reasons given. Table 7 shows the major reasons for relapse among participants who quit successfully at first but then relapsed at some point before filling out the survey (n=296). The most likely reasons for relapse were cravings for cigarettes (n=101, 34.1%) and alcohol consumption (n=91, 30.1%).

XSL•FO
RenderX

Table 6. Reasons for failure (n=112).

Reason for failure	Participants, n (%)
I was afraid of quitting	15 (42.9)
I did not believe that I could quit	11 (31.4)
I do not know	9 (25.7)
I rushed through the program and may have missed some concepts	7 (20)
I did not do all the steps of the final cigarette transformation ceremony	6 (17.1)
I smoked less than 10 cigarettes mindfully	6 (17.1)
I did not follow all the instructions	5 (14.3)
I took a break >2 days while doing the program	5 (14.3)
I did not believe in the content of the app	2 (5.1)
I did not like the content of the app	2 (5.1)

Table 7. Reasons for relapse (n=296).

Reason for relapse	Participants, n (%)
I still had bad cravings and I was unable to resist	101 (34.1)
I gave in while drinking alcohol	91 (30.1)
I faced a tragedy (eg, death of a loved one and bad breakup)	71 (24)
I became overconfident of my success	69 (23.3)
I felt self-destructive	47 (15.9)
I still believe smoking has some benefits	24 (8.1)
I did not do the program properly	18 (6.1)
My physical withdrawal symptoms were very bad	17 (5.7)
I gained a lot of weight	14 (4.7)
Other reasons (stress, peer pressure, etc)	27 (9.1)

Among those who relapsed or were unable to quit after completing the program (n=410), 80.7% (n=331) had a moderate to high motivation to quit at the time of submitting the survey. In total, 91% (n=377) said that they would consider using QuitSure for their next quitting attempt, while 46.1% (n=189) said that they will definitely use QuitSure to quit in the future.

Survey for Noncompleters of the Program

In total, 19,873 users had dropped off after starting the program and were sent the S-Incompleters survey, of which only 126 (0.6%) consented to participate in the study and submitted the survey.

Table 8 shows the reasons submitted for not completing the program (n=126). The most common reasons given for dropping off midway were a busy schedule (n=51, 40.5%), not enjoying the content of the program or having too much to read (n=25, 19.8%), and lack of belief that the program will work (n=20, 15.9%).

Table 8. Reasons for not completing the program (n=126).

Reason for not completing the program	Participants, n (%)
Busy	51 (40.5)
Did not enjoy content or too much reading	25 (19.8)
Lack of belief in the program	20 (15.9)
Quit smoking midway or cut down	16 (12.7)
Had to pay	14 (11.1)
Technical issues	11 (8.7)
Felt program was not working	9 (7.1)
Others	6 (4.7)
Was not ready	5 (4)

XSL•FO
RenderX

Discussion

Principal Findings

The purpose of the study was to understand whether the QuitSure program is an effective intervention for smoking cessation and can be implemented at scale to counteract the health and economic consequences of the tobacco epidemic. It was conducted via 2 web-based surveys. In total, 1299 participants submitted the S-Completers survey for program completers. A majority of 80.9% (1040/1286) maintained prolonged abstinence for 30 days after program completion, and 72.4% (725/1002) maintained 6-month prolonged abstinence after program completion. In total, 86.4% (770/891) of participants reported no severe withdrawal symptoms, while 41.9% (373/891) reported no withdrawal symptoms at all. Only 18.5% (165/891) experienced more than 5-kg weight gain after completing the program. Demographic variables such as gender and age did not significantly impact the program's success, but smoking quantity prior to doing the program and country of residence did have a significant impact on program efficacy. For those who relapsed, cravings and alcohol consumption were major factors, while program noncompletion was attributed to busy schedules or lack of belief in the program by the participants.

The program was able to achieve extended cessation for every category of smoker, from light to heavy with high efficacy rates, and low withdrawal symptoms after quitting. It is easy to navigate and uses simple language. The low price makes it affordable for smokers across most socioeconomic strata, and the easy-to-understand content makes it usable by anyone with a basic understanding of English. The fact that most participants who relapsed, or for whom the program did not work, continue believing in the program's potential is also a point in its favor. The difference in cessation rates in different countries indicates that the program requires some adaptations to be contextually and culturally relevant to the residents of certain countries.

When it comes to the feasibility of the program to be distributed to the population at large, the dropout rates of the program, at 59.4%, did not show improved program adherence and engagement compared to other health care apps [32]. While the very short 6-day length of the program likely increases completion rates, it requires approximately 1 hour of daily use. This high engagement requirement could be the reason why 40.5% (51/126) of participants who dropped off the program midway state being busy as the reason for noncompletion. The other major reason for dropping off the program was the length and style of the content. QuitSure could break down the program into a 30-day version with just 10-15 minutes of content per day for people who are busy or for whom the content seems too much. They could also add more graphics, videos, and design elements in the content to make it more appealing to users than plain, simple text. Lack of belief in the program's techniques was also shared as a reason for dropoff. The makers of the app can thus focus on informing the users about the scientific underpinnings of the techniques used in the program as well as include relevant references for their claims throughout the program.

The app is attempting to standardize and replicate an in-person deaddiction counseling program into a do-it-yourself app and uses many of these same psychological tools to achieve success for its users as in-person counseling [33-35]. The efficacy for those who completed the program and participated in the study seems high, indicating some degree of success. However, the program is of the do-it-yourself type and long enough that it requires high self-motivation and high intent to quit on the part of the user to complete the entire program. We do not have data on how many people dropped off even before starting because of the quantity of self-driven work required. A pre-post study analyzing dropout rates at every stage in the user journey will be required to evaluate the true feasibility of the app.

Limitations

The study had several limitations. A selection bias was created because the base sample selected was solely those who had already signed up for and completed the program, creating a closed cohort and a higher-than-normal intent to quit. Another limitation was the low response rate. Only 24.8% (1332/5365) of those who opened the email chose to submit the survey, allowing for a significant bias toward those for whom the program was successful. If we consider the program to have failed for all those who opened the email but did not submit the survey, the quit rate of QuitSure at the 30-day postprogram time point becomes just 19.4%. A recall bias may have resulted in false memories of withdrawal symptoms during the initial postprogram phase. The single measurement taken eliminates long-term cessation data of participants who only recently completed the program. Finally, the reward for filling out the survey may have motivated participants to give a false-positive response, based on an assumption that it would increase their likelihood of receiving the reward.

Ultimately, the obtained sample is not representative of the smoker population at large. To be able to understand the true feasibility and efficacy of the QuitSure program and counteract the above limitations, we would need to conduct a randomized controlled trial where the self-reported cessation of participants is confirmed via biochemical verification.

Comparison With Prior Work

Studies have shown that 46.3% of smokers who quit experience significant withdrawal symptoms ranging from anxiety, depression, irritability, and other physical symptoms [36,37]. The biggest strength of the QuitSure program is that only 13.6% (121/891) of the study participants faced severe withdrawal symptoms. Of the remaining participants, only around half faced even the milder withdrawal symptoms such as coughing and mild sleep disturbance, which are usually seen among all smokers upon quitting [38]. This could be a reflection of the program's focus on the psychological aspect of nicotine addiction via mindfulness, CBT, and reframing mental sets and beliefs, which have previously shown to reduce withdrawal symptoms after quitting [39,40]. Withdrawal symptoms are known to be a key contributor to relapse [36,37]. Therefore, it is likely that the increased effectiveness of the program and higher prolonged cessation rates are a result of these reduced withdrawal symptoms. However, QuitSure does not include any sort of NRT in its protocol. NRT is recommended by the

XSL•FO
RenderX

WHO, US Centers for Disease Control and Prevention, as well as the National Institute for Health and Care Excellence, United Kingdom [22,41,42] as an important complement to counseling and has been shown to significantly increase the success rates in psychology-based smoking cessation programs [34,43,44]. The hypothesis given is that NRT reduces withdrawal symptoms and craving levels. QuitSure could include a phased-out nicotine replacement regimen after the program to further increase its efficacy. This is especially important for heavy smokers, for whom the program was less effective.

The primary reason for relapse was due to still experiencing strong cravings for smoking. In fact, 75% (21/28) of participants who experienced greater than moderate levels of cravings after completing the program eventually relapsed. Currently, the QuitSure program does not address cravings management in any specific way after program completion, relying on the program itself to prevent the appearance of cravings at all. To address them and prevent their relapse, QuitSure can monitor craving levels after the program, with additional content and counseling for those who are struggling.

The second reason for relapse was alongside alcohol consumption. The app already recommends users not to drink any alcohol for the first week after quitting. It can extend this further and also give more guidelines on how to handle cravings when drinking alcohol.

Weight gain after quitting is another big concern among smokers, as there is evidence that nicotine reduces appetite, increases metabolism, and reduces food cravings [45,46]. Previous studies have shown that after quitting smoking, 35.4% of quitters had a weight gain over 5% of their body weight [47]. Of the study participants who were able to quit for even a brief period, 58.1% (518/891) had some weight gain. The QuitSure program should do more to specifically address the maladaptive thought patterns and beliefs connecting food, hunger, and smoking.

Previous studies have found that the higher an individual's app engagement is, the more they are likely to be able to quit smoking [19]. Thus, the QuitSure app needs to improve its engagement rates to increase program completion rates. Some tools the app developers can use to increase engagement that have previously demonstrated success are (1) gamification techniques like leaderboards, progress bars, and levels [19,48,49]; (2) small rewards to participants for every engagement milestone [48]; (3) personalizing notifications and reminders [49]; as well as (4) inclusion of a peer support group to improve program adherence and navigate postquitting withdrawal symptoms and cravings [50].

Overall, within the limitations of the study, the program shows high smoking cessation rates, low rates of withdrawal symptoms and cravings, and a generally positive experience for its users.

Conclusions

In total, 80.9% (1040/1286) of the survey respondents were able to achieve 30-day prolonged abstinence from smoking after program completion. The program also adheres to the WHO's 5As guideline for smoking cessation and includes psychological tools used in evidence-based in-person counseling protocols. However, there are many improvements in app engagement, program adherence, and postprogram support that can be made by the app developers. The high success rates, including prolonged cessation rates, among study participants are an indicator that QuitSure could be a useful tool for achieving smoking cessation at scale. Despite the severe limitations and selection biases of the study, the results make the QuitSure program a strong contender for further investigation. Health care institutions should consider and study the program's feasibility and efficacy in a more controlled setting.

Acknowledgments

The authors thank the users of the QuitSure app, whose participation and feedback played a pivotal role in this study. Their willingness to engage with the program and share their experiences provided valuable insights and enriched the findings.

Authors' Contributions

GMG and KB worked on the experimental design, data analysis, and paper writing. SM worked on data analysis and paper writing. All 3 authors reviewed the final paper.

Conflicts of Interest

GMG has no conflicts of interest. KB is a cofounder and equity holder in QuitSure. SM is an employee of QuitSure.

Multimedia Appendix 1
S-Completers survey questions.
[PDF File (Adobe PDF File), 42 KB - humanfactors_v11i1e49519_app1.pdf]

Multimedia Appendix 2
S-Incompleters survey questions.
[PDF File (Adobe PDF File), 41 KB - humanfactors_v11i1e49519_app2.pdf]

Multimedia Appendix 3

XSL•FO
RenderX

Participant funnel for the S-Completers survey.
[PDF File (Adobe PDF File), 43 KB - humanfactors_v11i1e49519_app3.pdf]

Multimedia Appendix 4
Global distribution of participants.
[PDF File (Adobe PDF File), 48 KB - humanfactors_v11i1e49519_app4.pdf]

References

1. GBD 2013 Risk Factors Collaborators, Forouzanfar MH, Alexander L, Anderson HR, Bachman VF, Biryukov S, et al. Global, regional, and national comparative risk assessment of 79 behavioural, environmental and occupational, and metabolic risks or clusters of risks in 188 countries, 1990-2013: a systematic analysis for the Global Burden of Disease Study 2013. Lancet 2015;386(10010):2287-2323. [doi: 10.1016/S0140-6736(15)00128-2] [Medline: 26364544]
2. Tobacco. World Health Organization. 2021. URL: https://www.who.int/news-room/fact-sheets/detail/tobacco [accessed 2023-04-03]
3. Economic trends in tobacco. Centers for Disease Control and Prevention. 2022. URL: https://www.cdc.gov/tobacco/data_statistics/fact_sheets/economics/econ_facts/index.htm [accessed 2023-03-29]
4. More than 100 reasons to quit tobacco. Pan American Health Organization. 2022. URL: https://www.paho.org/en/more-100-reasons-quit-tobacco [accessed 2023-03-29]
5. Alqahtani JS, Oyelade T, Aldhahir AM, Alghamdi SM, Almehmadi M, Alqahtani AS, et al. Prevalence, severity and mortality associated with COPD and smoking in patients with COVID-19: a rapid systematic review and meta-analysis. PLoS One 2020;15(5):e0233147 [FREE Full text] [doi: 10.1371/journal.pone.0233147] [Medline: 32392262]
6. Goodchild M, Nargis N, Tursan d'Espaignet E. Global economic cost of smoking-attributable diseases. Tob Control 2018 Jan;27(1):58-64 [FREE Full text] [doi: 10.1136/tobaccocontrol-2016-053305] [Medline: 28138063]
7. Babb S, Malarcher A, Schauer G, Asman K, Jamal A. Quitting smoking among adults—United States, 2000-2015. MMWR Morb Mortal Wkly Rep 2017 Jan 06;65(52):1457-1464 [FREE Full text] [doi: 10.15585/mmwr.mm6552a1] [Medline: 28056007]
8. Quitting smoking. National Cancer Institute. 2022. URL: https://progressreport.cancer.gov/prevention/quitting_smoking [accessed 2023-04-04]
9. Chaiton M, Diemert L, Cohen JE, Bondy SJ, Selby P, Philipneri A, et al. Estimating the number of quit attempts it takes to quit smoking successfully in a longitudinal cohort of smokers. BMJ Open 2016 Jun 09;6(6):e011045 [FREE Full text] [doi: 10.1136/bmjopen-2016-011045] [Medline: 27288378]
10. Lund M, Kvaavik E. Methods used in smoking cessation and reduction attempts: findings from help-seeking smokers. J Smok Cessat 2021;2021:6670628 [FREE Full text] [doi: 10.1155/2021/6670628] [Medline: 34306230]
11. Notley C, Gentry S, Livingstone-Banks J, Bauld L, Perera R, Hartmann-Boyce J. Incentives for smoking cessation. Cochrane Database Syst Rev 2019 Jul 17;7(7):CD004307 [FREE Full text] [doi: 10.1002/14651858.CD004307.pub6] [Medline: 31313293]
12. Fiore MC, Jaén CR, Baker TB, Bailey WC, Benowitz NL, Curry SJ, et al. Treating tobacco use and dependence: 2008 update. US Department of Health and Human Services, Public Health Service. 2008. URL: https://www.ahrq.gov/prevention/guidelines/tobacco/index.html [accessed 2024-03-23]
13. Etter JF, Stapleton JA. Nicotine replacement therapy for long-term smoking cessation: a meta-analysis. Tob Control 2006 Aug;15(4):280-285 [FREE Full text] [doi: 10.1136/tc.2005.015487] [Medline: 16885576]
14. Cahill K, Stevens S, Perera R, Lancaster T. Pharmacological interventions for smoking cessation: an overview and network meta-analysis. Cochrane Database Syst Rev 2013 May 31;2013(5):CD009329 [FREE Full text] [doi: 10.1002/14651858.CD009329.pub2] [Medline: 23728690]
15. Stitzer ML. Combined behavioral and pharmacological treatments for smoking cessation. Nicotine Tob Res 1999;1 Suppl 2:S181-S187; discussion S207. [doi: 10.1080/14622299050012041] [Medline: 11768178]
16. Kayhan Tetik B, Gedik Tekinemre I, Taş S. The effect of the COVID-19 pandemic on smoking cessation success. J Community Health 2021 Jun;46(3):471-475 [FREE Full text] [doi: 10.1007/s10900-020-00880-2] [Medline: 32643078]
17. Seo S, Cho SI, Yoon W, Lee CM. Classification of smoking cessation apps: quality review and content analysis. JMIR Mhealth Uhealth 2022 Feb 17;10(2):e17268 [FREE Full text] [doi: 10.2196/17268] [Medline: 35175213]
18. Zeng EY, Heffner JL, Copeland WK, Mull KE, Bricker JB. Get with the program: adherence to a smartphone app for smoking cessation. Addict Behav 2016 Dec;63:120-124 [FREE Full text] [doi: 10.1016/j.addbeh.2016.07.007] [Medline: 27454354]
19. Browne J, Halverson TF, Vilardaga R. Engagement with a digital therapeutic for smoking cessation designed for persons with psychiatric illness fully mediates smoking outcomes in a pilot randomized controlled trial. Transl Behav Med 2021 Sep 15;11(9):1717-1725 [FREE Full text] [doi: 10.1093/tbm/ibab100] [Medline: 34347865]

XSL•FO
RenderX

20. Abroms LC, Lee Westmaas J, Bontemps-Jones J, Ramani R, Mellerson J. A content analysis of popular smartphone apps for smoking cessation. Am J Prev Med 2013 Dec;45(6):732-736 [FREE Full text] [doi: 10.1016/j.amepre.2013.07.008] [Medline: 24237915]
21. Hoeppner BB, Hoeppner SS, Seaboyer L, Schick MR, Wu GWY, Bergman BG, et al. How smart are smartphone apps for smoking cessation? A content analysis. Nicotine Tob Res 2016 May;18(5):1025-1031 [FREE Full text] [doi: 10.1093/ntr/ntv117] [Medline: 26045249]
22. Toolkit for delivering the 5A's and 5R's brief tobacco interventions in primary care. World Health Organization. 2014. URL: https://iris.who.int/bitstream/handle/10665/112835/9789241506953_eng.pdf [accessed 2024-03-23]
23. Bold KW, Garrison KA, DeLucia A, Horvath M, Nguyen M, Camacho E, et al. Smartphone apps for smoking cessation: systematic framework for app review and analysis. J Med Internet Res 2023 Jul 13;25:e45183 [FREE Full text] [doi: 10.2196/45183] [Medline: 37440305]
24. Hersi M, Traversy G, Thombs BD, Beck A, Skidmore B, Groulx S, et al. Effectiveness of stop smoking interventions among adults: protocol for an overview of systematic reviews and an updated systematic review. Syst Rev 2019 Jan 19;8(1):28 [FREE Full text] [doi: 10.1186/s13643-018-0928-x] [Medline: 30660199]
25. Price rankings by country of cigarettes 20 pack (Marlboro) (markets). Numbeo. 2023. URL: https://www.numbeo.com/cost-of-living/country_price_rankings?itemId=17 [accessed 2023-04-10]
26. Darden ME, Hotchkiss JL, Pitts MM. Where there's smoke…: the wage impact of smoking. Federal Reserve Bank Atlanta's Policy Hub. 2020. URL: https://www.atlantafed.org/research/publications/policy-hub/2020/10/08/11-wage-impact-of-smoking.aspx [accessed 2024-03-19]
27. Eysenbach G. Improving the quality of web surveys: the Checklist for Reporting Results of Internet E-Surveys (CHERRIES). J Med Internet Res 2004 Sep 29;6(3):e34 [FREE Full text] [doi: 10.2196/jmir.6.3.e34] [Medline: 15471760]
28. What are good open rates, CTRs, and CTORs for email campaigns? Campaign Monitor. 2022. URL: https://www.campaignmonitor.com/resources/knowledge-base/what-are-good-email-metrics/#:~:text=According%20to%20our%202022%20Email,average%20between%2025%2D28%25.[accessed [accessed 2023-04-10]
29. Chung L. What is a good survey response rate for online customer surveys? Delighted. URL: https://delighted.com/blog/average-survey-response-rate [accessed 2023-04-10]
30. DeVita VT. How do I know if I'm a light, average, or heavy smoker? ShareCare. 2023. URL: https://www.sharecare.com/health/quit-smoking/how-do-i-know-if-im- [accessed 2023-04-12]
31. Japuntich SJ, Leventhal AM, Piper ME, Bolt DM, Roberts LJ, Fiore MC, et al. Smoker characteristics and smoking-cessation milestones. Am J Prev Med 2011 Mar;40(3):286-294 [FREE Full text] [doi: 10.1016/j.amepre.2010.11.016] [Medline: 21335259]
32. Torous J, Lipschitz J, Ng M, Firth J. Dropout rates in clinical trials of smartphone apps for depressive symptoms: a systematic review and meta-analysis. J Affect Disord 2020 Feb 15;263:413-419. [doi: 10.1016/j.jad.2019.11.167] [Medline: 31969272]
33. Manjula DC, Shekar V, Reddy RCJ. Effect of behavioural counselling in tobacco cessation. J Sci Dent 2018;8(2):58-60 [FREE Full text] [doi: 10.5005/jsd-8-2-58]
34. Lindson-Hawley N, Thompson TP, Begh R. Motivational interviewing for smoking cessation. Cochrane Database Syst Rev 2015 Mar 02(3):CD006936. [doi: 10.1002/14651858.CD006936.pub3] [Medline: 25726920]
35. Blebil AQ, Sulaiman SAS, Hassali MA, Dujaili JA, Zin AM. Impact of additional counselling sessions through phone calls on smoking cessation outcomes among smokers in Penang State, Malaysia. BMC Public Health 2014 May 16;14:460 [FREE Full text] [doi: 10.1186/1471-2458-14-460] [Medline: 24886549]
36. McLaughlin I, Dani JA, De Biasi M. Nicotine withdrawal. Curr Top Behav Neurosci 2015;24:99-123 [FREE Full text] [doi: 10.1007/978-3-319-13482-6_4] [Medline: 25638335]
37. Lancaster T, Stead LF. Individual behavioural counselling for smoking cessation. Cochrane Database Syst Rev 2017 Mar 31;3(3):CD001292 [FREE Full text] [doi: 10.1002/14651858.CD001292.pub3] [Medline: 28361496]
38. Thornley S, McRobbie H, Lin RB, Bullen C, Hajek P, Laugesen M, et al. A single-blind, randomized, crossover trial of the effects of a nicotine pouch on the relief of tobacco withdrawal symptoms and user satisfaction. Nicotine Tob Res 2009 Jun;11(6):715-721. [doi: 10.1093/ntr/ntp054] [Medline: 19454549]
39. Cui ZY, Li YH, Liu Z, Li L, Nie XQ, Zhou XM, et al. The experience of tobacco withdrawal symptoms among current smokers and ex-smokers in the general population: findings from nationwide China Health Literacy Survey during 2018-19. Front Psychiatry 2022;13:1023756 [FREE Full text] [doi: 10.3389/fpsyt.2022.1023756] [Medline: 36713929]
40. Mhende J, Bell SA, Cottrell-Daniels C, Luong J, Streiff M, Dannenfelser M, et al. Mobile delivery of mindfulness-based smoking cessation treatment among low-income adults during the COVID-19 pandemic: pilot randomized controlled trial. JMIR Form Res 2021 Jul 23;5(7):e25926 [FREE Full text] [doi: 10.2196/25926] [Medline: 34033580]
41. Quit smoking. Centers for Disease Control and Prevention. 2023. URL: https://www.cdc.gov/tobacco/quit_smoking/index.htm [accessed 2023-04-13]
42. Recommendations on promoting quitting. National Institute for Health and Care Excellence. 2023. URL: https://www.nice.org.uk/guidance/ng209/chapter/Recommendations-on-promoting-quitting [accessed 2023-04-13]

XSL•FO
RenderX

43. Jódar-Sánchez F, Carrasco Hernández L, Núñez-Benjumea FJ, Mesa González MA, Moreno Conde J, Parra Calderón CL, et al. Using the Social-Local-Mobile app for smoking cessation in the SmokeFreeBrain project: protocol for a randomized controlled trial. JMIR Res Protoc 2018 Dec 06;7(12):e12464 [FREE Full text] [doi: 10.2196/12464] [Medline: 30522992]
44. Carrasco-Hernandez L, Jódar-Sánchez F, Núñez-Benjumea F, Moreno Conde J, Mesa González M, Civit-Balcells A, et al. A mobile health solution complementing psychopharmacology-supported smoking cessation: randomized controlled trial. JMIR Mhealth Uhealth 2020 Apr 27;8(4):e17530 [FREE Full text] [doi: 10.2196/17530] [Medline: 32338624]
45. Chen H, Hansen MJ, Jones JE, Vlahos R, Bozinovski S, Anderson GP, et al. Cigarette smoke exposure reprograms the hypothalamic neuropeptide Y axis to promote weight loss. Am J Respir Crit Care Med 2006 Jun 01;173(11):1248-1254. [doi: 10.1164/rccm.200506-977OC] [Medline: 16531608]
46. Bloom EL, Farris SG, DiBello AM, Abrantes AM. Smoking-related weight and appetite concerns and use of electronic cigarettes among daily cigarette smokers. Psychol Health Med 2019 Feb;24(2):221-228 [FREE Full text] [doi: 10.1080/13548506.2018.1537495] [Medline: 30346797]
47. Jeremias-Martins E, Chatkin JM. Does everyone who quit smoking gain weight? A real-world prospective cohort study. J Bras Pneumol 2019 Feb 25;45(1):e20180010 [FREE Full text] [doi: 10.1590/1806-3713/e20180010] [Medline: 30810645]
48. Garaialde D, Cox AL, Cowan BR. Designing gamified rewards to encourage repeated app selection: effect of reward placement. Int J Hum-Comput Stud 2021 Sep;153:102661 [FREE Full text] [doi: 10.1016/j.ijhcs.2021.102661]
49. Jakobsen PR, Christensen JR, Nielsen JB, Søndergaard J, Jarbøl DE, Olsen MH, et al. Identification of important factors affecting use of digital individualised coaching and treatment of type 2 diabetes in general practice: a qualitative feasibility study. Int J Environ Res Public Health 2021;18(8):3924 [FREE Full text] [doi: 10.3390/ijerph18083924] [Medline: 33917999]
50. Cheung YTD, Chan CHH, Lai CKJ, Chan WFV, Wang MP, Li HCW, et al. Using WhatsApp and Facebook online social groups for smoking relapse prevention for recent quitters: a pilot pragmatic cluster randomized controlled trial. J Med Internet Res 2015 Oct 22;17(10):e238 [FREE Full text] [doi: 10.2196/jmir.4829] [Medline: 26494159]

Abbreviations

CBT: cognitive behavioral therapy
CHERRIES: Checklist for Reporting Results of Internet E-Surveys
HIPAA: Health Insurance Portability and Accountability Act
NRT: nicotine replacement therapy
WHO: World Health Organization

Edited by Y Senathirajah; submitted 31.05.23; peer-reviewed by D Ford, L Huang; comments to author 02.09.23; revised version received 28.11.23; accepted 28.02.24; published 06.05.24.

Please cite as:
Goldgof GM, Mishra S, Bajaj K
Efficacy of the QuitSure App for Smoking Cessation in Adult Smokers: Cross-Sectional Web Survey
JMIR Hum Factors 2024;11:e49519
URL: https://humanfactors.jmir.org/2024/1/e49519
doi:10.2196/49519
PMID:38709553

©Gregory M Goldgof, Shweta Mishra, Kriti Bajaj. Originally published in JMIR Human Factors (https://humanfactors.jmir.org), 06.05.2024. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on https://humanfactors.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Digital Adherence Technologies Linked to Mobile Money Incentives for Medication Adherence Among People Living With Tuberculosis: Mixed Methods Feasibility and Acceptability Study

Angella Musiimenta[1,2], PhD; Wilson Tumuhimbise[1,2], PhD; Esther Atukunda[3], PhD; Aaron Mugaba[1,2], MBA; Sebastian Linnemayr[4], PhD; Jessica Haberer[5,6], MD

[1]Faculty of Computing and Informatics, Mbarara University of Science and Technology, Mbarara, Uganda
[2]Angels Compassion Research and Development Initiative, Mbarara, Uganda
[3]Faculty of Medicine, Mbarara University of Science and Technology, Mbarara, Uganda
[4]Rand Corporation, Santa Monica, CA, United States
[5]Harvard Medical School, Boston, MA, United States
[6]Massachusetts General Hospital Center for Global Health, Boston, MA, United States

Corresponding Author:
Angella Musiimenta, PhD
Faculty of Computing and Informatics
Mbarara University of Science and Technology
Plot 8-19 Mbarara Kabale Road
Mbarara, P.O. Box 653
Uganda
Phone: 256 776820598
Email: amusiimenta@must.ac.ug

Abstract

Background: Complementing digital adherence technologies (DATs) with mobile money incentives may improve their utility in supporting tuberculosis medication adherence, yet the feasibility and acceptability of this integrated approach remain unclear.

Objective: This study aims to describe the feasibility and acceptability of a novel DAT intervention called My Mobile Wallet composed of real-time adherence monitoring, SMS text message reminders, and mobile money incentives for tuberculosis medication adherence in a low-income setting.

Methods: We purposively recruited people living with tuberculosis from the Mbarara Regional Referral Hospital in Mbarara, Uganda, who (1) were starting tuberculosis treatment at enrollment or within the past 4 weeks, (2) owned a mobile phone, (3) were able to use SMS test messaging, (4) were aged ≥18 years, and (5) were living in Mbarara district. At study exit (month 6), we used interviews and questionnaires informed by the unified theory of acceptance and use of technology (UTAUT) to collect feasibility and acceptability data, reflecting patients' experiences of using each component of My Mobile Wallet. Feasibility also included tracking the functionality of the adherence monitor (ie, an electronic pillbox) as well as SMS text message and mobile money delivery. We used a content analytical approach to inductively analyze qualitative data and Stata (version 13; StataCorp LLC) to analyze quantitative data.

Results: All 39 participants reported that the intervention was feasible because it was easy for them to use (eg, access and read SMS text messages) and worked as expected. Almost all SMS text messages (6880/7064, 97.4%) were sent as planned. The transmission of adherence data from the monitor worked well, with 98.37% (5682/5776) of the data transmitted as planned. All participants additionally reported that the intervention was acceptable because it helped them take their tuberculosis medication as prescribed; the mobile money incentives relieved them of tuberculosis-related financial burdens; SMS text message reminders and electronic pillbox–based alarms reminded them to take their medication on time; and participants perceived real-time adherence monitoring as "being watched" while taking their medication, which encouraged them to take their medication on time to demonstrate their commitment. The intervention was perceived as a sign of care, which eventually created emotional support and a sense of connectedness to health care. Participants preferred daily SMS text message reminders (32/39, 82%) to reminders linked to missed doses (7/39, 18%), citing the fact that tuberculosis medication is taken daily.

XSL•FO
RenderX

Conclusions: The use of real-time adherence monitoring linked to SMS text message reminders and mobile money incentives for tuberculosis medication adherence was feasible and acceptable in a low-resource setting where poverty-based structural barriers heavily constrain tuberculosis treatment and care.

(JMIR Hum Factors 2024;11:e47996) doi:10.2196/47996

KEYWORDS

digital adherence technologies; real-time monitoring; SMS text message reminders; mobile money; financial incentives; tuberculosis; medication adherence; user-centered approach

Introduction

Background

Tuberculosis treatment adherence remains challenging in Uganda. Constraints to tuberculosis medication adherence include a lack of transport to the clinic to pick up the drugs and forgetfulness [1]. Digital adherence technologies (DATs) are being explored to encourage adherence to tuberculosis medication [2,3]. Recently, we showed that real-time adherence monitors linked to SMS text message reminders were potentially useful in reminding patients to take their medication and encouraging tuberculosis medication adherence in rural Uganda [4]. However, tuberculosis is well known to be a disease of poverty [5], and the lack of money may potentially limit the usefulness of DATs (eg, the inability to afford transport to pick up the medications) [6]. Although effective tuberculosis treatment has existed since the 1940s and is available for free, many people delay seeking treatment, struggle with medication adherence, or do not complete their treatment because of poverty [1]. This is because tuberculosis leads to the loss of productivity of patients and their caregivers, resulting in additional costs for patients in the form of transport to and from the clinic and may lead to loss of employment for fear of spreading the disease to other people [7]. Currently, in Uganda, 53% of patients living with tuberculosis take loans or sell property to meet the costs of their tuberculosis care [8]. Interventions are necessary to overcome the poverty-based structural barriers to tuberculosis treatment, including unconditional transport to and from the clinic. According to the End TB strategy of the World Health Organization (WHO), the use of social protection schemes (such as transport to the clinic and meals) could lower the financial burden of tuberculosis [9]. A recent systematic literature review and meta-analysis by Richterman et al [10] defines cash transfers as cash payments provided to specific beneficiaries. The review indicates that cash transfer interventions may improve treatment success among patients with pulmonary tuberculosis, although the review expresses the need for more research regarding the effectiveness of sensitive cash transfers for tuberculosis care, especially in low-income countries [10].

The use of mobile money technology (money sent, received, or saved on mobile phones) is a promising tool for delivering health-related cash transfers; for instance, mobile money enabled pregnant women to save for maternal health care in Kenya [11], while a progressive incentive scheme to reward private physicians and community health care workers enhanced identification and referral of suspected tuberculosis cases and treatment tracking in Pakistan [12]. The use of mobile money transfers to incentivize patients living with tuberculosis to take their drugs may potentially improve their adherence to medication [13]. However, the use of mobile money services in the context of health care is still in its infancy, and the limited research in this area reports mixed results [14].

My Mobile Wallet

My Mobile Wallet is a DAT intervention composed of a real-time adherence monitor, SMS text message reminders, and mobile money incentives (known as WiseCash). The financial incentives are meant to motivate participants to take their medication as well as enable them to attend their clinic appointments for pill refills. The intervention was developed through user-centered approaches [15], and we previously published formative qualitative findings indicating the anticipated benefits and challenges of using the intervention for tuberculosis medication adherence in rural Uganda [13]. In brief, participants reported that the intervention could remind them to take their medication as well as support, and motivate tuberculosis medication adherence. However, they expressed concerns about the possible unintended tuberculosis status disclosure as well as the possibility of using the money for other competing demands. This information was then used to refine and improve My Mobile Wallet.

This paper presents the feasibility and acceptability of a pilot study implementing My Mobile Wallet. Specifically, we present the practical experiences of people living with drug-sensitive tuberculosis who used the intervention during their 6-month tuberculosis treatment period.

Methods

Study Design and Setting

This study used a convergent mixed methods study design. The study recruited people living with tuberculosis from the tuberculosis clinic at the Mbarara Regional Referral Hospital (MRRH) in Mbarara in southwestern Uganda. The tuberculosis clinic provides care to an estimated 600 people living with tuberculosis annually. All newly diagnosed people living with tuberculosis receive free tuberculosis medication and are counseled about the benefits of tuberculosis medication at the tuberculosis clinic. At the MRRH, the recommended directly observed therapy approach (which advises that patients should take their medication as they are watched by a health care provider or treatment supporter) is not used for monitoring medication adherence due to the costs involved for both people living with tuberculosis and the health care workers. Instead, people living with tuberculosis are treated with the 2HRZE regimen (isoniazid, rifampin, pyrazinamide, and ethambutol for 2 months) in the initiation phase, with clinic visits every 2

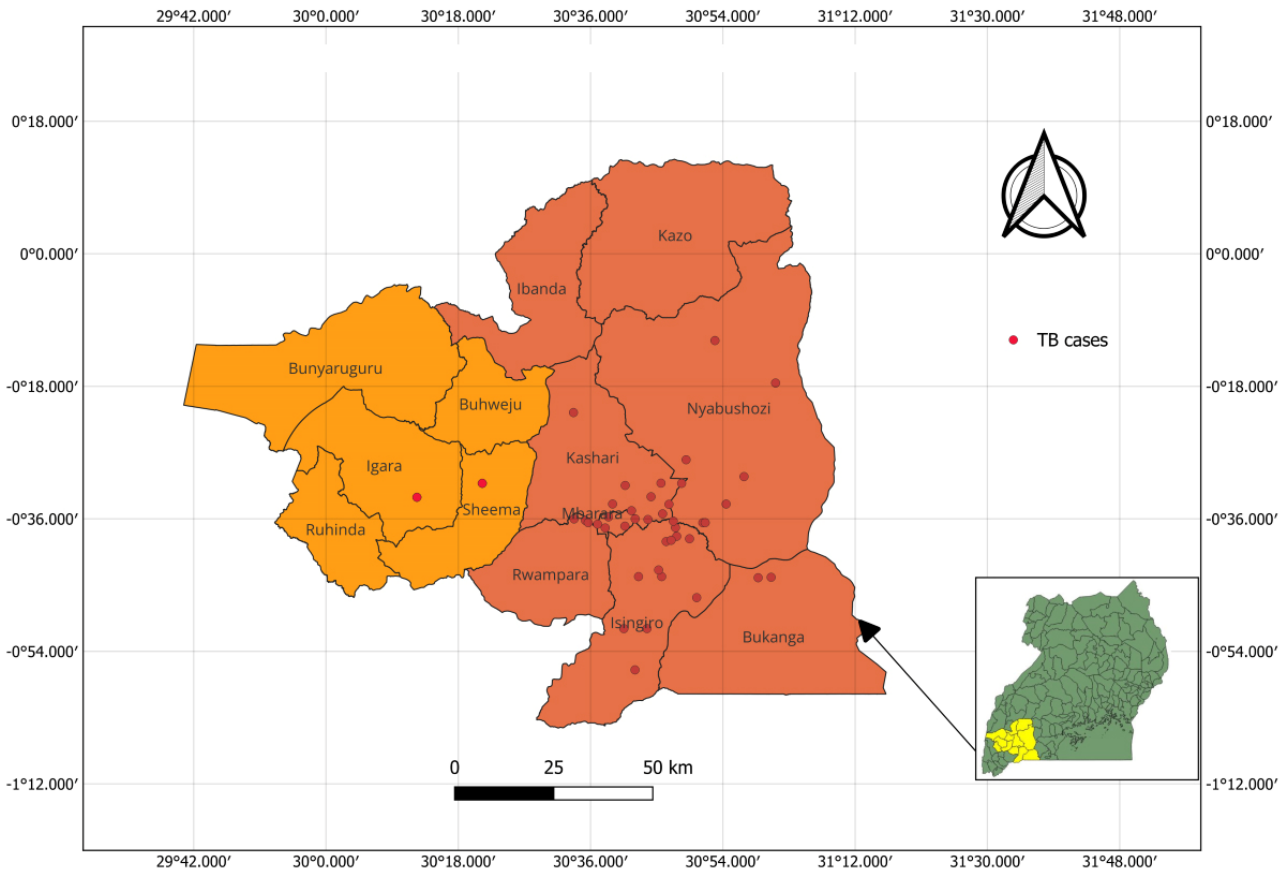
XSL•FO
RenderX

weeks. At the end of the 2-month period, they return to the tuberculosis clinic for a sputum conversion check. Those who become smear negative continue with the 4HR regimen (isoniazid plus rifampin for 4 months) in the continuation phase, with monthly clinic visits. Those with positive test results receive GeneXpert to exclude rifampicin resistance; subsequent treatment is then individualized. Treatment may be extended up to a full year to compensate for missed medication pick-ups or doses.

Selection of Study Participants

Between July 2022 and October 2022, we recruited participants at the MRRH according to the following inclusion criteria: (1) newly diagnosed with tuberculosis per the clinic records and starting tuberculosis treatment at enrollment or within the past 4 weeks, (2) owning a mobile phone, and (3) living in Mbarara district (Figure 1). We excluded individuals who were unwilling or unable to provide informed consent due to severe mental conditions per the clinical records and those unable to use mobile money–based SMS text messaging (we trained potential participants and tested this skill at recruitment). We purposively sampled patients to achieve relatively balanced representation by HIV status and sex to solicit diverse perspectives.

Figure 1. The study area map and geographic distribution of participants. TB: tuberculosis.

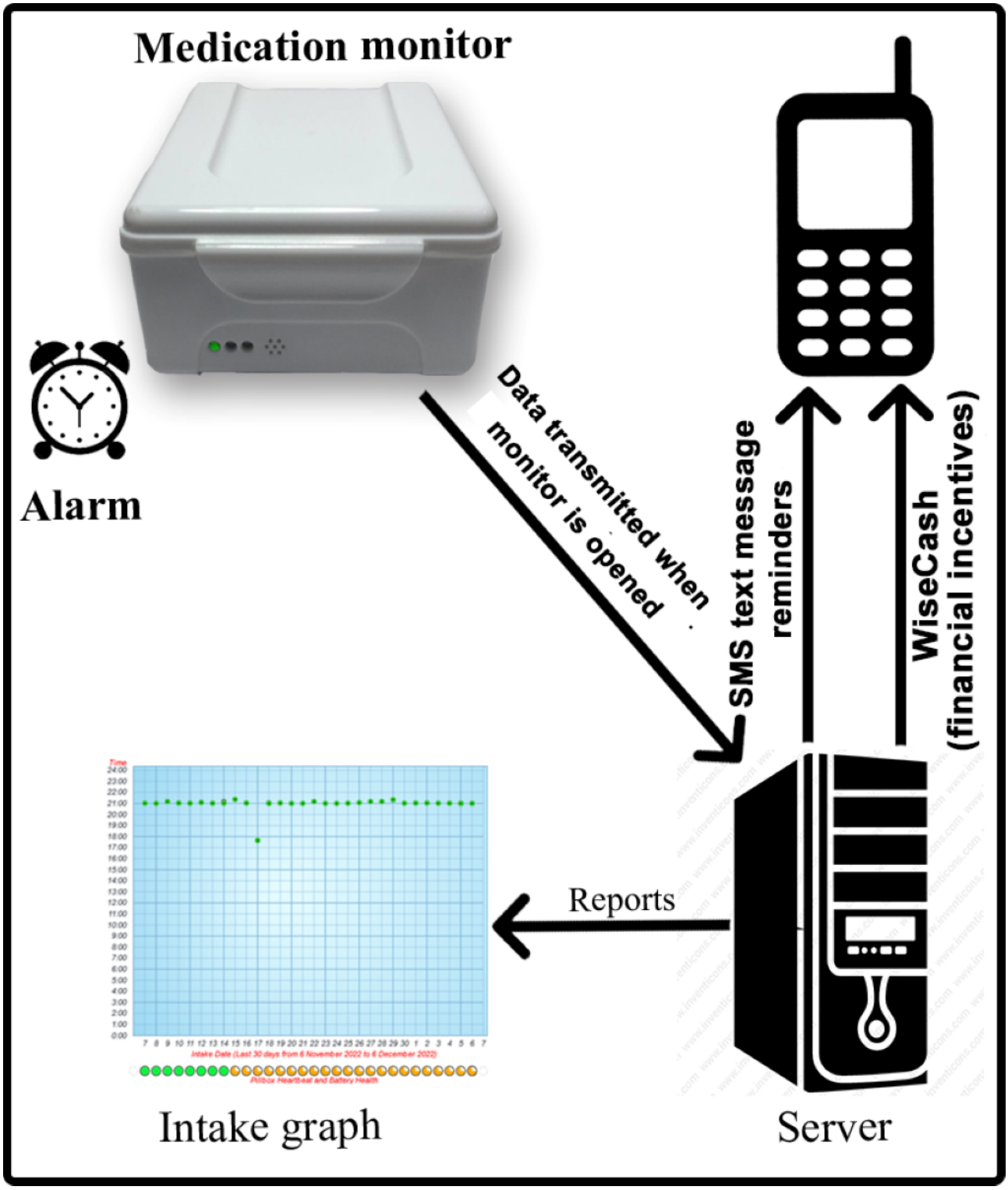


Intervention Technology

Details of the My Mobile Wallet intervention are described elsewhere [13]. Briefly, as shown in Figure 2, the intervention is composed of the following 3 components: a real-time medication monitor (Wisepill evriMED1000) to monitor medication adherence by sending signals when opened (the monitor records a date-and-time stamp as a proxy for taking medication, and it has an option to set an alarm sound to remind patients to take their medication); SMS text message reminders sent to users' mobile phones to remind them to take their medication as prescribed (reminders are sent daily for 2 months, after which they are triggered as needed by missed or delayed doses); and the WiseCash app, which uses a tailored mobile money platform for sending financial incentives for transport to the clinic and motivating medication adherence.

XSL•FO
RenderX

Figure 2. The My Mobile Wallet intervention diagram.



Study Procedures

We first oriented each participant to the My Mobile Wallet intervention components. We explained and demonstrated how the real-time adherence monitor (Wisepill) works, including how it monitors medication adherence and sends a signal to researchers every time it is opened, how the monitor makes an alarm sound to remind patients to take their medication, and how to open and close the monitor to put in or retrieve medication. Participants were then asked to demonstrate how the monitor works. Next, we explained to them how the intervention sends daily SMS text message reminders (30 minutes before medication-taking time) to remind participants to take their medication for the first 2 months. We also explained how the intervention sends SMS text message reminders for the next 4 months only if the monitor is not opened within an hour of the expected time (known as triggered SMS text message reminders). We then explained how the intervention transfers USh 28,000 (approximately equivalent to US $8; we decided upon this amount based on the transport costs, which had

XSL•FO
RenderX

increased during the COVID-19 pandemic and had not reduced at the time of implementing this intervention) as an unconditional monthly mobile money incentive to mobile phones belonging to people living with tuberculosis for facilitating transport to the clinic (Figure 1 shows a visual representation of the geographic distribution of the participants) from the date of study recruitment until the end of their 6-month treatment period; furthermore, the participants were informed that USh 5250 (approximately equivalent to US $1.50) would be transferred as a monthly conditional medication adherence incentive to those with a medication adherence rate of ≥90% as ascertained from the real-time adherence monitor. The transfer of the transport incentives required patients to inform the research staff about their next date of appointment so that it could be input into the WiseCash application to allow automatic triggering of the transfer of the transport incentive a day before their next visit. We decided upon the medication adherence rate of >90% because evidence shows that adherence below this level does not yield favorable treatment outcomes [16].

Data Collection

We administered a baseline demographic and sociobehavioral questionnaire to participants at enrollment, which included age, sex, tuberculosis medication specifications (drugs and planned dosing times), and mobile phone number and use. We used the interviewer-administered approach for administering questionnaires to elicit quantitative data orally from the participants (ie, closed-ended questions read out in the participants' local language by the researcher, with participants answering the questions orally). Several validated surveys were adopted and included in this questionnaire (eg, the Duke-UNC Functional Social Support Questionnaire for measuring social support [17], the asset index scale to assess socioeconomic status [18], the depression section of the Hopkins Symptom Checklist for assessing depression [19], the Household Food Insecurity Access Scale for measuring food insecurity [20], the Alcohol Use Disorders Identification Test for assessing alcohol consumption [21], and the Internalized AIDS-Related Stigma Scale for assessing stigma [22]). Feasibility was ascertained by tracking the functionality of the monitor and SMS text message and mobile money delivery. The unified theory of acceptance and use of technology (UTAUT) model [23], given its track record of predicting a substantial portion of the acceptance of digital health interventions, provided a basis for developing surveys and interview guides for capturing participants' views on feasibility and acceptability at study exit (month 6). The UTAUT model asserts that the adoption of technology is influenced by four major constructs as perceived by an individual user: (1) performance expectancy or perceived usefulness of the intervention (in this case, My Mobile Wallet) (2) effort expectancy or perceived ease of use of the intervention, (3) social norms (how others perceive the individual's use of the intervention), and (4) facilitating conditions (the availability of technical and organizational infrastructure to support the use of the intervention). A structured exit questionnaire aimed at eliciting closed-ended information from participants regarding their experiences of using My Mobile Wallet was administered. This was a Likert scale questionnaire that sought to explore the extent to which participants liked or disliked the functionalities of the intervention. Qualitative open-ended interviews, by contrast, elicited in-depth information about participants' experiences using each component of My Mobile Wallet (the real-time monitor, SMS text message reminders and monitor alarms, and the WiseCash application), including benefits and challenges related to the technologies. Authors WT and ATM (who are trained in qualitative research and research ethics) conducted the semistructured in-depth interviews with participants in a private space at a research office near the MRRH until thematic saturation was reached at the 30th participant interview. Each interview lasted between 30 and 60 minutes and was conducted in the local language (Runyankole), digitally recorded, transcribed, and translated into English. After each interview, author AM reviewed the transcripts for quality, clarity, and detail.

Data Analysis

We followed the UTAUT model [23] to review transcripts for content related to acceptability. We then developed a coding scheme, used it to code the data, and reviewed the coded data to develop descriptive categories. We mapped the descriptive categories onto the domains of the UTAUT model's four major constructs that influence technology adoption: (1) performance expectancy or perceived usefulness, (2) effort expectancy or perceived ease of use, (3) social norms, and (4) facilitating conditions. Illustrative quotations were then selected from the coded data. After the completion of the codebook, we applied the codes using NVivo 11 (Lumivero). We followed the COREQ (Consolidated Criteria for Reporting Qualitative Research) [24] checklist in reporting qualitative results. Feasibility metrics and the quantitative assessment of acceptability were analyzed descriptively by WT and ATM using Stata 13.

Ethical Considerations

The institutional review committees of Mbarara University of Science & Technology (MUST-2021-102) and the Uganda National Council for Science and Technology (HS1688ES) approved this study. All participants provided signed informed consent before study participation.

Results

Demographic Characteristics

Of the 54 screened participants, we excluded 5 (9%) for not owning a mobile phone, 5 (9%) for not living within 60 kilometers of Mbarara district, and 4 (7%) for having mobile phone numbers that were not registered for mobile money service. Thus, 40 (74%) of the 54 screened participants were enrolled in the study and used the intervention for 6 months. Of these 40 participants, 1 (2%) was lost to follow-up (her mobile phone was unreachable). As indicated in Table 1, of the 40 participants, 24 (60%) were female, 27 (68%) had coinfection with HIV, 34 (85%) had no regular or fixed income, 18 (45%) did not study beyond primary level (typically attended by children aged 6-12 years), 40 (100%) perceived their social support to be insufficient, and 21 (53%) reported severe food insecurity. The participants' median age was 38 (IQR 28-54) years, and, before joining the study, they were on medication for a median of 4 (IQR 2.5-8) weeks.

XSL•FO
RenderX

Table 1. Baseline demographic characteristics of participants (n=40).

Characteristics	Values
Age (y), median (IQR)	38 (28-54)
Weeks on medication before joining the study, median (IQR)	4 (2.5-8)
Sex, n (%)	
Male	24 (60)
Female	16 (40)
Education, n (%)	
None	3 (8)
P1-P7[a]	18 (45)
Ordinary level	9 (22)
Advanced level	3 (8)
Tertiary level	7 (18)
Income (fixed wages or salary), n (%)	
Yes	6 (15)
No	34 (85)
Heavy alcohol consumption, n (%)	
Yes	1 (2)
No	39 (98)
Enough social support (no), n (%)	40 (100)
Food insecurity, n (%)	
Yes	21 (52)
No	19 (48)
Probable depression, n (%)	
Yes	1 (2)
No	39 (98)
Asset index scale[b], n (%)	
Lowest quartile	16 (40)
25%-100% quartiles	24 (60)
HIV status, n (%)	
Negative	13 (32)
Positive	27 (68)

[a]In the Ugandan education system, primary school (P1-P7) is often attended by children aged 6 to 12 years.

[b]Index was measured using the measure proposed by Filmer and Pritchett [18].

As indicated in Table 2, at baseline, half of the participants (20/40, 50%) did not share their mobile phones with anyone, 85% (34/40) checked their SMS text messages more frequently than *often* in a week, 75% (30/40) *often* used mobile money, 88% (35/40) preferred receiving SMS text message reminders daily because medication taking is a daily activity, 68% (27/40) preferred SMS text message reminders that are not easily related to tuberculosis (eg, “Hello today”) to avoid unwanted tuberculosis status disclosure, and 38% (15/40) preferred receiving mobile money incentives for transport to the clinic a day before the clinic visit to avoid using the money for other competing needs.

XSL•FO
RenderX

Table 2. Mobile phone use and intervention preferences at baseline (n=40).

Questions	Participants, n (%)
Who else uses your mobile phone?	
Spouse	3 (8)
Family member	16 (40)
Neighbor	1 (2)
No one else	20 (50)
Check for SMS text messages in a week	
Never	1 (2)
Less than *often*	5 (12)
More than *often*	34 (85)
Use of mobile money	
Less than *often*	10 (25)
More than *often*	30 (75)
Reasons for delay in checking for SMS text messages last week[a]	
Mobile phone not charged	20 (50)
Mobile phone was used by someone else	2 (5)
No adequate signal	9 (22)
Mobile phone not functioning	1 (2)
Used by someone else	2 (5)
Preferred frequency of receiving SMS text message reminders	
Daily	35 (88)
Weekly	5 (12)
Preferred content for SMS text message reminders	
Not easily related to TB[b] (eg “Hello today”)	27 (68)
Easily related to TB (eg, “Take your TB drugs”)	13 (32)
SMS text message language preference	
Local language	22 (55)
English	18 (45)
When to send the mobile money incentive for transport to the clinic	
1 day before the clinic visit	25 (63)
2 days before the clinic visit	10 (25)
>2 days before the clinic visit	5 (12)

[a]Reasons for delay in checking for SMS text messages last week, n=21, 52%.

[b]TB: tuberculosis.

Exit Survey Results

All participants self-reported that it was easy for them to access and read the mobile money SMS text messages as well as the medication-taking reminders (39/39, 100%) and open the Wisepill device to retrieve their medication (39/39, 100%). In addition, all participants (39/39, 100%) received the mobile money incentive for transport to the clinic as expected, received the medication adherence incentives as expected, and reported that the real-time adherence monitor worked as expected. All participants (39/39, 100%) additionally reported that the mobile money incentives, the Wisepill device, and the SMS text message reminders helped them take their tuberculosis medication on time or as prescribed.

The average adherence rate ascertained from the real-time monitors was 90.4% (SD 8.6%), and 24 (60%) of the 40 participants had an adherence rate of >90%.

As indicated in Table 3, almost all participants (38/40, 95%) opted to be reminded by both SMS text message reminders and alarms from the real-time monitor. Of the 40 participants, 2 (5%) requested study staff to switch off the alarms on their

XSL•FO
RenderX

monitors at enrollment because they anticipated being inconvenienced by the sound. Participants preferred daily SMS text message reminders (32/39, 82%) to reminders linked to missed doses (7/39, 18%), citing the fact that tuberculosis medication is taken daily. All participants reported that the mobile money incentives were sent as expected. However, 4 (10%) of the 40 participants received cash (once during the study period) as refund for the transport fare instead of being sent the mobile money incentive for transport to the clinic. These participants did not inform the study staff on time about their next clinic visit date, which had to be input into the WiseCash application to allow automatic triggering of the transfer of the incentive a day before the clinic visit. Almost all the SMS text messages (6880/7064, 97.4%) were sent as planned. The transmission of adherence data from the monitor worked well, with 98.37% (5682/5776) of the data transmitted as planned. No real-time adherence monitor malfunctioned during the study period.

Table 3. Feasibility and acceptability of the My Mobile Wallet intervention.

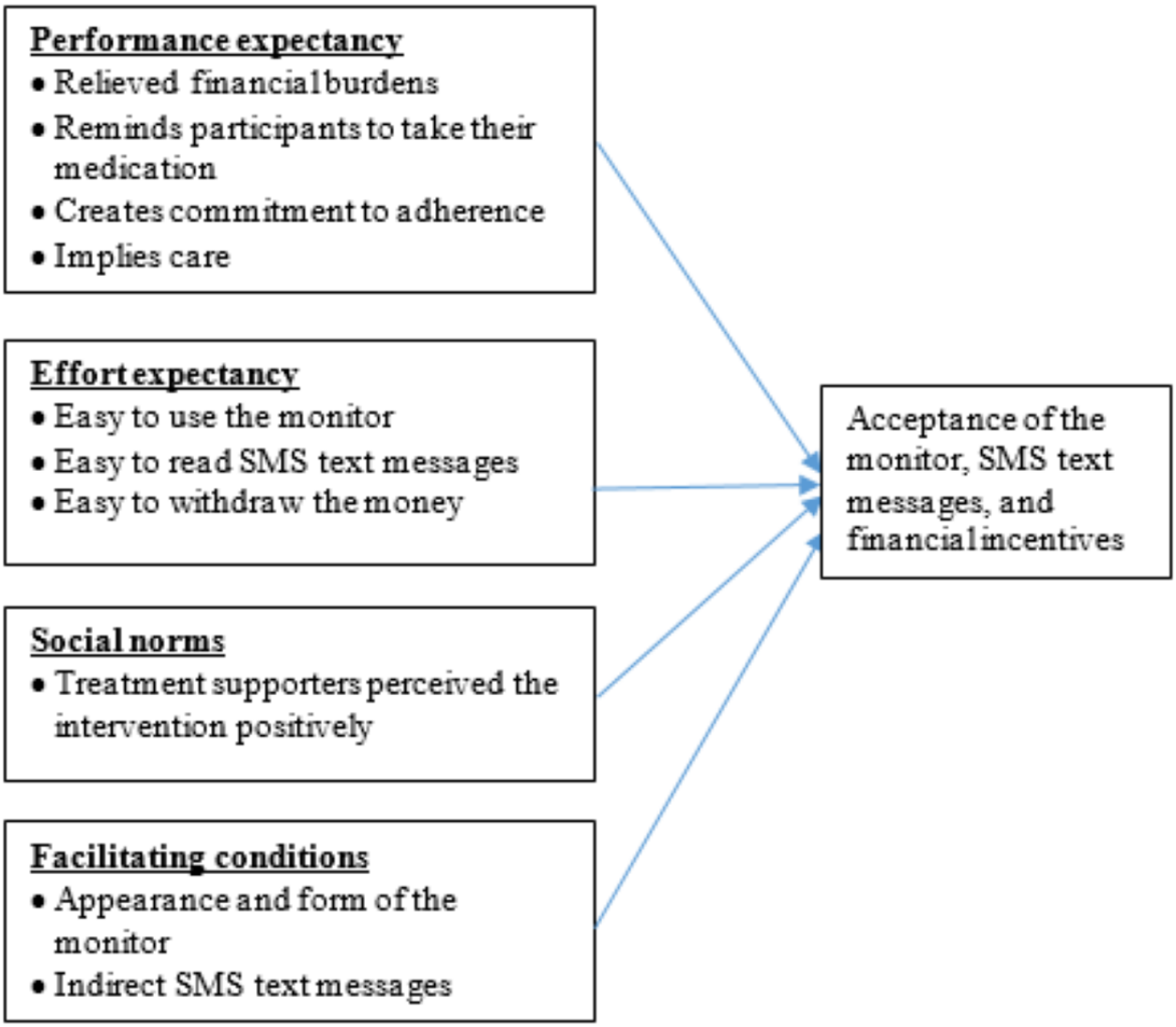
Feasibility and acceptability of SMS text messages	Values
Preference of SMS text message reminders versus device-based alarms (recorded at recruitment; n=40), n (%)	
Participants who opted to be reminded by SMS text message reminders only	0 (0)
Participants who opted to be reminded by Wisepill device alarm only	2 (5)
Participants who opted for SMS text message reminders plus Wisepill device alarm	38 (95)
SMS text message reminders (automatically ascertained from the intervention; N=7064), n (%)	
Total number of SMS text message reminders sent	6880 (97.4)
SMS text message reminders not sent due to technical challenges (eg, poor network)	184 (2.6)
Mobile money incentives (automatically ascertained from the intervention; N=40), n (%)	
Transport incentives not sent	0 (0)
Transport incentives sent unnecessarily	0 (0)
Adherence incentives not sent	0 (0)
Adherence incentives sent unnecessarily	0 (0)
Feasibility and acceptability of the real-time adherence monitor, n (%)	
Data loss due to technical issues with the real-time monitors (days when the monitor was not opened; automatically ascertained from the intervention; N=5776)	94 (1.63)
Device malfunction (N=40; devices that malfunctioned and were replaced)	0 (0)
Devices successfully returned by participants (N=40; the device used by the participant who was lost to follow-up was later recovered from her treatment supporter)	40 (100)

Interview Results: Intervention Acceptability

Acceptability is presented following the constructs of the UTAUT model (Figure 3) of the performance expectancy, effort expectancy, social norms, and facilitating conditions associated with the intervention.

XSL•FO
RenderX

Figure 3. Organization of the qualitative data on acceptability following the unified theory of acceptance and use of technology (UTAUT) model.



Performance Expectancy or Perceived Usefulness

Mobile Money Incentives Relieve Participants of Tuberculosis-Related Financial Burdens

Before being enrolled in the study, some of the participants struggled to meet the basic costs of their tuberculosis care, including transport to the tuberculosis clinic and the cost of food and drinks needed to take the medication. These were mainly participants with no regular source of income, as well as those staying far from the tuberculosis clinic. They described relying on begging and borrowing money to meet their tuberculosis treatment costs. Unfortunate instances where begging and attempts to borrow money were not successful resulted in missed clinic appointments for pill refills due to lack of transport, which consequently resulted in missed medication. Others reported not taking their medication due to lack of food and drinks because they feared becoming weaker after taking their medication on an empty stomach. Participants reported that the mobile money transport incentives enabled them to meet the cost of transport to the clinic (eg, for pill refills) as well as the costs of meeting the basic tuberculosis treatment needs, such as food and drinks for taking their medication, thereby relieving them (as their treatment supporters) of the financial burdens associated with tuberculosis treatment and care. One participant stated as follows:

> *Before you started sending me mobile money, there were times when I would request people to borrow me money or help me with transport to the hospital, but sometimes they would also not be in a position to give me money, so I would miss picking my medication from the hospital...I stopped working when I got sick. I stay far from the hospital.This study helped a lot by sending me [money for] transport to the hospital so that I don't miss taking my medication on time.* [Male patient, aged 61 years]

Participants narrated how medication adherence incentives encouraged them to take their medication on time to meet the monthly target of ≥90% medication adherence, which in turn helped them meet the cost of the basic food and drinks they needed for taking medication:

> *I knew that I would be given money after getting ≥90% medication adherence, so I made sure that I was taking my medication well in order to be sent the*

XSL•FO
RenderX

money. This disease made me too weak to work; yet, I needed money to buy food and porridge. [Female patient, aged 51 years]

Whenever I would receive a message about my adherence percentage, and it's below, it would motivate me to be serious so that next time I don't miss out again. [Male patient, aged 39 years]

SMS Text Message Reminders and the Real-Time Adherence Monitor Enabled Participants to Take Their Medication on Time

Participants reported that the SMS text message reminders and monitor-based alarms enabled them to take their medication on time. These technologies served as medication reminders, thereby addressing forgetfulness, which was common in participants who were not yet used to taking medication regularly. They were also useful for busy participants who could easily forget their medication-taking time due to other competing demands on their time. Participants reported being able to make the necessary preparations (eg, obtaining food and drinks and going back home in case the medication had been left behind at home) after receiving the daily SMS text message reminders (which were sent 30 minutes before their medication-taking time), thereby enabling them to take their medication on time:

The SMS [text message] reminders were very helpful because I was still learning how to take medicine in time, so they helped me in getting used to medication taking because they were coming every day, so my body eventually got used to the time. [Female patient, aged 35 years]

At times, I would get too busy at my video library and forget taking my medication, but whenever I would receive the message, I would close the business immediately, go home, eat some food, prepare a drink, and use it to take my medication on time. [Male patient, aged 27 years]

Using a Real-Time Adherence Monitor Creates Commitment to Medication Adherence

Participants perceived real-time adherence monitoring as "being watched" while taking their medication. This perception was welcomed and encouraged them to take their medication on time to demonstrate their commitment to the health care providers who they felt were concerned about their health and would not be happy with nonadherence:

When I started using the device, I felt touched knowing that there are people who are concerned about my life to the extent of using the device to watch me take medication yet they are not even my relatives or friends or people I knew before. This gave me morale to swallow my medication to play my part especially because they would be seeing whether or not I am taking my medication and they would probably feel bad if I miss taking [it]. [Female patient, aged 40 years]

Whenever I felt like not taking the medication, I always got motivated me to take medication because I knew that you people cared for me so much to the extent that you gave me this monitor and kept texting me to remind me to take medication and even sent me money to go to the clinic. I felt encouraged because you were really interested in seeing my health condition improve. [Female patient, aged 26 years]

Although monitoring created commitment to medication adherence, it is noteworthy that the primary motivation for taking medication on time as reported by participants was the need to recover their good health and live longer. A participant stated as follows:

Whenever I saw it [the device], I knew it was going to report me, so I chose to take the commitment of swallowing the medication. But, the main issue was, I really needed to recover from this disease because I loved my life and wanted to save it by getting well as soon as possible. [Male patient, aged 57 years]

Receiving Financial Incentives and Reminders and Being Monitored Implies Care

Overview

Receiving the mobile money incentives and SMS text message or alarm reminders and being monitored via the real-time adherence monitor was perceived by participants as signs that the health care providers cared about them, which eventually created emotional support and a sense of connectedness that countered depressive feelings. A participant describes how she changed her mind about committing suicide as a result of using the technologies:

I was about to stop taking the medication and die because I developed self-rejection. I was in pain, and I had no one to help me, but you people encouraged me to take the medication when you put me in this study and started sending me texts, alarms, and mobile money to support me to take my medication. I dropped the idea of suicide because you people cared for me and loved me even more than I loved myself. Thank you for saving my life because I would be dead by now. [Female patient, aged 33 years]

Effort Expectancy or Perceived Ease of Use: The Intervention Was Easy to Use

After the participants' initial orientation to using the real-time adherence monitor, they found it easy to use for taking their medication. Participants, including a few (3/40, 8%) who never went to school, reported finding it easy to read the SMS text messages sent to them. In addition, they reported that it was easy to withdraw money from mobile money agents because the agents are readily available:

It was very easy to use the container [the real-time adherence monitor]; you open it the same way a food box is [opened], put your medication [in], and start using it; that is all. I did not have to charge it or do any other thing with it. [Male patient, aged 42 years]

Although I did not go to school, I can read messages written in my local language, so, reading the messages on [the mobile] phone was not a problem at all. [Female patient, aged 28 years]

XSL•FO
RenderX

> *There are so many mobile money agents around. It was easy for me to withdraw my money from them.* [Male patient, aged 35 years]

Social Norms or Other People's Perceptions of the Intervention

Positive Perceptions From Treatment Supporters

Participants reported that their treatment supporters approve of them using the intervention to support their medication adherence to get well. In addition, participants stated that because of the financial incentives, their treatment supporters were relieved of the financial burden of having to take care of the financial needs of the participants:

> *When my wife saw the container and the messages, she was happy knowing that I was being supported by the hospital; she believed the support would help [me] get well soon. It was also a relief for her when you sent me money; I stopped working when I got sick, and before your assistance, I was only relying on getting money from her small shop for transport to the hospital and getting other basic needs like food.* [Male patient, aged 32 years]

Possibility of Inappropriate Use of the Financial Incentives

One participant reported how her husband initially misappropriated the financial incentives intended to pay for her transport to the clinic. Although the participant did not miss visiting the clinic, she had to keep begging her husband for money for transport to the clinic. Sometimes she would have to walk part of the distance due to insufficient transport funds provided by her husband:

> *My husband never wants me to own any money and always forces [me] to give him my money. So, whenever you would send me money on the [mobile] phone, he would force me to give him the whole of it. I would then have to go through the hassle of begging him to give me the money for my hospital visit, and sometimes the money he would give me for transport would not be enough.... But after giving you my new SIM card [details], which he did not know [about], I started receiving and using the money for transport to the hospital.* [Female patient, aged 33 years]

Facilitating Conditions: Appearance and Form of the Monitor

Participants reported that they liked the appearance and form of the real-time adherence monitor. Specifically, they liked the monitor's design, which resembled a food box, and its size, which they thought was reasonable because it accommodated all their pills; the absence of tuberculosis-related labels that could link them to the disease; and the hard outer cover that kept their medicines safe and clean, all of which motivated them to use the monitor:

> *The container looks like a food bowl, so people can easily think you have carrying some food in it; it is also big enough to carry all my medicine, and has no any TB-related word.* [Female patient, aged 35 years]

Indirect SMS Text Messages

To avoid unwanted status disclosure, participants preferred SMS text messages that could not easily link them to tuberculosis:

> *I chose the message "come and eat" because for me I knew what it reminded me to eat, but other people even if they saw it on my [mobile] phone would not know what I was going to eat.* [Male patient, aged 38 years]

Discussion

Principal Findings

Drawing on the UTAUT model, this paper describes the feasibility and acceptability of My Mobile Wallet, a DAT intervention composed of a real-time adherence monitor, SMS text message reminders, and mobile money incentives (WiseCash) for tuberculosis medication adherence in rural southwestern Uganda. Generally, we found that the intervention was technically feasible because it functioned as expected. All participants reported that it was easy for them to use the intervention; they could access and read the mobile money SMS text messages as well as the medication reminders, and they were able to open the Wisepill device to retrieve their medication. Participants reported receiving the mobile money incentives for transport to the clinic and the medication adherence incentives as expected. The SMS text messages and real-time adherence monitor also worked as expected: the SMS text messages were sent as planned, the transmission of adherence data from the monitor worked well, and no monitor malfunctioned for the entire period of the study.

Concerning acceptability, participants reported being relieved of tuberculosis-related financial burdens as a result of receiving the mobile money incentives. SMS text message reminders and real-time monitor-based alarms reminded participants to take their medication on time. Daily SMS text message reminders were preferred to reminders triggered by missed doses. Participants' preference for daily SMS text message reminders even in the treatment continuation phase (from month 4 onward) was surprising because one would assume that during this phase, they were nearly getting used to taking their medication and therefore did not require to be reminded daily. Patients' preference for daily SMS text message reminders (for taking their medications) over weekly SMS text message reminders was also reported in our previous tuberculosis study [4] and HIV study [25]. As tuberculosis medications are taken daily, daily SMS text message reminders are preferred because they are aligned with the medication-taking frequency.

Participants perceived real-time adherence monitoring as "being watched" while taking their medication, which was welcomed and encouraged them to take their medication on time to demonstrate their commitment. Receiving the mobile money incentives and SMS text message or alarm reminders and being monitored via the real-time adherence monitor was perceived by participants as signs that the health care providers cared about them. Their experiences with the intervention eventually

XSL•FO
RenderX

created emotional support and a sense of connectedness that countered depressive feelings among the participants. Inappropriate use of the mobile money transport incentives was reported only rarely.

Limitations

The main limitation of this study is the possibility of social desirability bias in the data collected from interviews and surveys. The mobile money incentives in particular may have influenced participants, given the prevalence of poverty-based structural barriers to tuberculosis treatment in Uganda, a low-resource setting.

Comparison With Prior Work

We are not aware of any study that reports the impact of a DAT intervention composed of a real-time adherence monitor, SMS text message reminders, and mobile money incentives on tuberculosis medication adherence. However, it should be noted that some studies using some components of this intervention exist; for instance, a recent systematic review and meta-analysis on cash interventions to improve tuberculosis outcomes concluded that these interventions could improve tuberculosis treatment success and completion among patients in low- and middle-income countries [10]; in this review, only 1 randomized control trial in Peru [26] was identified, and the authors of the review noted that the evidence is still weak. In addition, the use of face-to-face cash transfers or transport vouchers as incentives has been reported to be acceptable in facilitating adherence to tuberculosis diagnostic evaluation in Uganda [27]. Furthermore, receiving monthly financial incentives face-to-face enabled patients living with tuberculosis in Nigeria to purchase food and get transport to the clinic [28], while, in Uganda, receiving a one-time cash transfer upon sputum submission supported tuberculosis testing completion among patients [29]. Although the receipt of unconditional cash transfers through a direct benefit scheme supported registered patients living with tuberculosis to meet their nutrition requirements in India, the scheme had no significant effect on treatment outcome [30]. In our study, individuals with no regular source of income and those living far from the clinic benefited most from the mobile money incentives; participants used the transport incentives to cover the cost of transport to the clinic, while they used the financial incentives conditional on high medication adherence to buy food and drinks required to take their medication. This approach could potentially address the financial insecurities that continue to constrain medication adherence [4]. In Uganda, 53% of the patients living with tuberculosis take loans or sell property to meet the costs of their tuberculosis care [8]. Although there is limited research in this area, our study indicates that mobile money incentives can potentially relieve the financial burden that tuberculosis places not only on patients but also on their treatment supporters. An incentive as small as US $1 can increase the tuberculosis cure rate and reduce treatment loss to follow-up in Uganda [31]. Although this study estimated an average transport cost of US $8 (based on the COVID-19–pandemic-induced transport cost increases) and provided the same amount for transport to all participants, using GPS information to estimate and provide transport costs according to each participant's distance from their home to the clinic could be a better option.

The reported practice of a husband taking the mobile money transport incentive from the wife shows the effects of poverty as well as the complexity of implementing mobile money incentives in low-resource settings and cultures where some people still believe in male-exclusive ownership of resources [32]. This scenario could result in an inappropriate use of the incentives, thus limiting the impact of the intervention. An inclusive approach that engages men in the implementation of such an intervention (eg, through awareness creation) might mitigate the risk of the incentives being misappropriated.

This is the first study to report on the feasibility of real-time monitoring linked to SMS text message reminders and mobile money incentives for tuberculosis medication adherence. In the same setting, we had previously reported that using SMS text message reminders linked to real-time monitoring is feasible and acceptable for supporting tuberculosis medication adherence [6]. This study provides insights regarding the integration of financial incentives with these technologies to support access to tuberculosis medication from the hospital and motivate medication adherence.

Participants' medication adherence ascertained from the real-time adherence monitor was quite high. The receipt of financial incentives that was conditional upon a particular adherence target (≥90%) resulted in participants taking their medication on time in order to hit the target for financial incentives. In addition, participants' awareness of the fact that their medication adherence was being watched or monitored through the real-time monitor motivated them to take their medication well in order not to disappoint those monitoring them. Notably, the reported adherence was ascertained from the monitor in the form of the monitor being opened, which was used as a proxy for medication taking. Overall, the real-time monitoring approach can potentially be more reliable than participant self-reports, which are highly subject to social desirability bias. However, although it was not reported in our findings, instances of opening the monitor without taking medication (such as accidental openings or opening the monitor to increase the chances of getting incentives) may constrain the feasibility of the intervention. In the ongoing phase of the study, we are supplementing the real-time adherence monitoring with hair analysis (assessing tuberculosis drug levels in participants' hair) to improve objectivity.

Although a few SMS text messages (184/7064, 2.6%) and some adherence data (94/5776, 1.63%) could not be sent by the SMS text messaging application and the real-time adherence monitor, respectively, mainly due to technical issues such as poor network, the intervention was otherwise feasible. The feasibility of this intervention could be attributed to the rapid evolution and adoption of mobile phone technologies in Uganda, including among populations based in rural areas and considered economically marginalized [33]. The applications for the SMS text message reminders and mobile money incentives were tailored from the existing mobile phone infrastructure, which likely facilitated use by participants who were already familiar using SMS text messaging and mobile money services in their

XSL•FO
RenderX

regular routines. By leveraging the existing mobile phone infrastructure, these technologies can potentially bridge the current gaps in access to health care services between economically advantaged populations and populations considered disadvantaged, consequently contributing to equitable access to health care. In addition, the fact that all participants owned personal mobile phones, had the ability to read SMS text message reminders and mobile money SMS text messages, and had reliable mobile network (per the enrolment criteria) could have contributed to the feasibility of the intervention. Different feasibility results may be yielded if this intervention is implemented in populations with fewer resources or in privileged populations.

Concerning acceptability, participants perceived the intervention's functionalities of sending timely medication-taking reminders (through SMS text messages and monitor-based alarms), financially supporting medication adherence (through mobile money incentives), and monitoring medication taking (through the adherence monitor) as supportive and taking care of them. For the participants, this perception created a sense of connectedness with health care providers and countered depressive feelings after their tuberculosis diagnosis. It also encouraged them to adhere to taking their medication as a way of appreciating the care and proving their commitment to taking an active role in their own health with the ultimate goal of regaining their health. Such emotional support can also potentially empower patients to cope with the stigma and discrimination that are often associated with tuberculosis [34]. Importantly, there is evidence that emotional and social support can improve tuberculosis treatment success rates [35]. The use of various tuberculosis medication adherence technologies (including SMS text messages and real-time adherence monitors) was perceived by participants to reduce visits to clinics and increase access to social supporters in a variety of settings [36]. In South Africa, the use of a real-time adherence monitor (Wisepill evriMED1000) was acceptable for prompting a stepwise differentiated care approach for tuberculosis medication adherence, composed of SMS text messages, telephone calls, home visits, and motivational counseling, in response to missed doses ascertained from the monitor [37]. Other studies referencing patient experiences of using real-time adherence monitoring linked to SMS text message reminders for antiretroviral adherence support among people living with HIV in Uganda also reported perceptions of being cared for as a result of using the technologies [25]. Although there are differences between HIV and tuberculosis, they are both diseases of poverty, and result in stigma, and discrimination. The findings regarding the adherence monitor and the SMS text message aspects of the intervention were indeed similar in this and another [25] study, indicating the strong potential of the intervention in this and potentially other similar settings.

Conclusions

In sum, we found the My Mobile Wallet intervention (composed of real-time adherence monitoring linked to SMS text message reminders and mobile money incentives) for tuberculosis medication adherence to be feasible and acceptable in a low-resource setting where poverty-based structural barriers heavily constrain tuberculosis treatment and care. The intervention worked as expected, and participants found it easy to use. The intervention relieved participants of the burden of tuberculosis treatment costs, reminded them to take their medication on time, and provided emotional support that made them feel connected to care.

On the basis of the findings from this study, we are now planning a randomized controlled trial (registered on ClinicalTrials.gov; NCT05656287) for assessing the full-scale feasibility, acceptability, and impact of My Mobile Wallet.

Acknowledgments

The authors would like to acknowledge the contributions of Wisepill Technologies, Yo Uganda Limited (Yo!), and the study participants. The study was funded by a grant from the Eunice Kennedy Shriver National Institute of Child Health and Human Development of the National Institutes of Health (R21HD107985). JH is supported through a grant from the National Institute of Mental Health (K24MH114732). EA is supported by the Fogarty International Center of the National Institutes of Health (K43TWO11004).

Data Availability

The data sets generated and analyzed during this study are available from the corresponding author on reasonable request.

Conflicts of Interest

None declared.

References

1. Yoeli E, Rathauser J, Rand D. The resurgence of tuberculosis is behavioral, not medical. Nudges can fix it. STAT. URL: https://www.statnews.com/2019/10/25/tuberculosis-behavior-text-based-nudges/ [accessed 2024-03-03]
2. Boutilier JJ, Yoeli E, Rathauser J, Owiti P, Subbaraman R, Jónasson JO. Can digital adherence technologies reduce inequity in tuberculosis treatment success? Evidence from a randomised controlled trial. BMJ Glob Health 2022 Dec 01;7(12):e010512 [FREE Full text] [doi: 10.1136/bmjgh-2022-010512] [Medline: 36455988]

XSL•FO
RenderX

3. Rao JS, Diwan V, Kumar AA, Varghese SS, Sharma U, Purohit M, et al. Acceptability of video observed treatment vs. directly observed treatment for tuberculosis: a comparative analysis between south and central India. Wellcome Open Res 2022 May 13;7:152. [doi: 10.12688/wellcomeopenres.17865.1]
4. Musiimenta A, Tumuhimbise W, Mugaba AT, Muzoora C, Armstrong-Hough M, Bangsberg D, et al. Digital monitoring technologies could enhance tuberculosis medication adherence in Uganda: mixed methods study. J Clin Tuberc Other Mycobact Dis 2019 Dec;17:100119 [FREE Full text] [doi: 10.1016/j.jctube.2019.100119] [Medline: 31788561]
5. Barter DM, Agboola SO, Murray MB, Bärnighausen T. Tuberculosis and poverty: the contribution of patient costs in sub-Saharan Africa--a systematic review. BMC Public Health 2012 Nov 14;12(1):980 [FREE Full text] [doi: 10.1186/1471-2458-12-980] [Medline: 23150901]
6. Musiimenta A, Tumuhimbise W, Atukunda EC, Mugaba AT, Musinguzi N, Muzoora C, et al. The feasibility, acceptability, and preliminary impact of real-time monitors and SMS on tuberculosis medication adherence in southwestern Uganda: findings from a mixed methods pilot randomized controlled trial. PLOS Glob Public Health 2023 Dec 5;3(12):e0001813 [FREE Full text] [doi: 10.1371/journal.pgph.0001813] [Medline: 38051699]
7. The Uganda tuberculosis communication strategy. Ministry of Health, Uganda. 2010. URL: http://library.health.go.ug/sites/default/files/resources/The%20Uganda%20Tuberculosis%20Communication%20Strategy.pdf [accessed 2024-04-29]
8. Direct and indirect costs due to tuberculosis and proportion of tuberculosis-affected households experiencing catastrophic costs due to TB in Uganda. Ministry of Health, Uganda. 2019. URL: https://www.health.go.ug/cause/direct-and-indirect-costs-due-to-tuberculosis-and-proportion-of-tuberculosis-affected-households-experiencing-catastrophic-costs-due-to-tb-in-uganda-febrauary-2019/ [accessed 2024-04-29]
9. Implementing the end TB strategy: the essentials (No. WHO/HTM/TB/2015.31). World Health Organization. 2015. URL: https://www.who.int/publications/i/item/9789240065093 [accessed 2024-04-29]
10. Richterman A, Steer-Massaro J, Jarolimova J, Luong Nguyen LB, Werdenberg J, Ivers LC. Cash interventions to improve clinical outcomes for pulmonary tuberculosis: systematic review and meta-analysis. Bull World Health Organ 2018 Jul 01;96(7):471-483 [FREE Full text] [doi: 10.2471/BLT.18.208959] [Medline: 29962550]
11. van Bastelaer T, Woodman B, Chatterji M, Long D. Saving for delivery through technology: an inquiry into an electronic commitment savings product for maternal healthcare in Kenya. Int J Health Plann Manage 2016 Oct 28;31(4):e254-e272. [doi: 10.1002/hpm.2292] [Medline: 25920603]
12. Mobile money for health case study compendium. United States Agency for International Development. 2015. URL: https://www.hfgproject.org/wp-content/uploads/2015/10/HFG-Mobile-Money-Compendium_October-2015.pdf [accessed 2024-04-29]
13. Musiimenta A, Tumuhimbise W, Atukunda E, Mugaba A, Linnemayr S, Haberer J. Digital adherence technologies and mobile money incentives for management of tuberculosis medication among people living with tuberculosis: mixed methods formative study. JMIR Form Res 2023 Apr 12;7:e45301 [FREE Full text] [doi: 10.2196/45301] [Medline: 37043263]
14. Lacroze E, Bärnighausen T, De Neve JW, Vollmer S, Ratsimbazafy RM, Emmrich PM, et al. The 4MOTHERS trial of the impact of a mobile money-based intervention on maternal and neonatal health outcomes in Madagascar: study protocol of a cluster-randomized hybrid effectiveness-implementation trial. Trials 2021 Oct 21;22(1):725 [FREE Full text] [doi: 10.1186/s13063-021-05694-8] [Medline: 34674741]
15. Brown T. Change by Design: How Design Thinking Transforms Organizations and Inspires Innovation. New York, NY: HarperCollins; 2019.
16. Imperial MZ, Nahid P, Phillips PP, Davies GR, Fielding K, Hanna D, et al. A patient-level pooled analysis of treatment-shortening regimens for drug-susceptible pulmonary tuberculosis. Nat Med 2018 Nov 5;24(11):1708-1715 [FREE Full text] [doi: 10.1038/s41591-018-0224-2] [Medline: 30397355]
17. Broadhead WE, Gehlbach SH, de Gruy FV, Kaplan BH. The Duke-UNC functional social support questionnaire: measurement of social support in family medicine patients. Med Care 1988 Jul;26(7):709-723. [doi: 10.1097/00005650-198807000-00006] [Medline: 3393031]
18. Filmer D, Pritchett LH. Estimating wealth effects without expenditure data-or tears: an application to educational enrollments in states of India. Demography 2001 Feb;38(1):115. [doi: 10.2307/3088292]
19. Bolton P, Wilk CM, Ndogoni L. Assessment of depression prevalence in rural Uganda using symptom and function criteria. Soc Psychiatry Psychiatr Epidemiol 2004 Jun;39(6):442-447. [doi: 10.1007/s00127-004-0763-3] [Medline: 15205728]
20. Coates J, Swindale A, Bilinsky P. Household food insecurity access scale (HFIAS) for measurement of food access: indicator guide: version 3. FANTA Project. 2007. URL: https://www.fantaproject.org/sites/default/files/resources/HFIAS_ENG_v3_Aug07.pdf [accessed 2024-04-29]
21. Bush K, Kivlahan DR, McDonell MB, Fihn SD, Bradley KA. The AUDIT alcohol consumption questions (AUDIT-C): an effective brief screening test for problem drinking. Ambulatory care quality improvement project (ACQUIP). Alcohol use disorders identification test. Arch Intern Med 1998 Sep 14;158(16):1789-1795. [doi: 10.1001/archinte.158.16.1789] [Medline: 9738608]
22. Kalichman SC, Simbayi LC, Cloete A, Mthembu PP, Mkhonta RN, Ginindza T. Measuring AIDS stigmas in people living with HIV/AIDS: the Internalized AIDS-related stigma scale. AIDS Care 2009 Jan 11;21(1):87-93. [doi: 10.1080/09540120802032627] [Medline: 19085224]

XSL•FO
RenderX

23. Venkatesh V, Morris MG, Davis GB, Davis FD. User acceptance of information technology: toward a unified view. MIS Q 2003;27(3):425-478. [doi: 10.2307/30036540]
24. Tong A, Sainsbury P, Craig J. Consolidated criteria for reporting qualitative research (COREQ): a 32-item checklist for interviews and focus groups. Int J Qual Health Care 2007 Dec 16;19(6):349-357. [doi: 10.1093/intqhc/mzm042] [Medline: 17872937]
25. Ware NC, Pisarski EE, Tam M, Wyatt MA, Atukunda E, Musiimenta A, et al. The Meanings in the messages: how SMS reminders and real-time adherence monitoring improve antiretroviral therapy adherence in rural Uganda. AIDS 2016 May 15;30(8):1287-1294 [FREE Full text] [doi: 10.1097/QAD.0000000000001035] [Medline: 26807967]
26. Wingfield T, Tovar MA, Huff D, Boccia D, Montoya R, Ramos E, et al. Socioeconomic support to improve initiation of tuberculosis preventive therapy and increase tuberculosis treatment success in Peru: a household-randomised, controlled evaluation. Lancet 2017 Feb;389:S16. [doi: 10.1016/s0140-6736(17)30412-9]
27. Kadota JL, Nabwire S, Nalugwa T, White JS, Cattamanchi A, Katamba A, et al. Patient perspectives and willingness to accept incentives for tuberculosis diagnostic evaluation in Uganda. Value Health Reg Issues 2021 Sep;25:48-56 [FREE Full text] [doi: 10.1016/j.vhri.2020.12.005] [Medline: 33773327]
28. Ukwaja KN, Alobu I, Mustapha G, Onazi O, Oshi DC. 'Sustaining the DOTS': stakeholders' experience of a social protection intervention for TB in Nigeria. Int Health 2017 Mar 01;9(2):112-117. [doi: 10.1093/inthealth/ihx001] [Medline: 28204499]
29. Shete PB, Kadota JL, Nanyunja G, Namale C, Nalugwa T, Oyuku D, et al. Evaluating the impact of cash transfers on tuberculosis (ExaCT TB): a stepped wedge cluster randomised controlled trial. ERJ Open Res 2023 May 01;9(3):00182-2023 [FREE Full text] [doi: 10.1183/23120541.00182-2023] [Medline: 37342088]
30. Dave JD, Rupani MP. Does direct benefit transfer improve outcomes among people with tuberculosis? - A mixed-methods study on the need for a review of the cash transfer policy in India. Int J Health Policy Manag 2022 Dec 06;11(11):2552-2562 [FREE Full text] [doi: 10.34172/ijhpm.2022.5784] [Medline: 35174676]
31. Baluku JB, Nakazibwe B, Twinomugisha B, Najjuuko R, Isabella N, Nassozi S, et al. One dollar incentive improves tuberculosis treatment outcomes in programmatic settings in rural Uganda. Sci Rep 2021 Sep 29;11(1):19346 [FREE Full text] [doi: 10.1038/s41598-021-98770-7] [Medline: 34588552]
32. Kane EW. Men's and women's beliefs about gender inequality: family ties, dependence, and agreement. Sociol Forum 1998 Dec;13(4):611-637.
33. The Collaboration on International ICT Policy for East and Southern Africa. National information technology survey report 2017/2018. National Information Technology Authority-Uganda. 2018. URL: https://s3-eu-west-1.amazonaws.com/s3.sourceafrica.net/documents/120158/National-IT-Survey-April-10th.pdf [accessed 2024-04-29]
34. Musiimenta A, Tumuhimbise W, Atukunda EC, Mugaba AT, Muzoora C, Armstrong-Hough M, et al. Mobile health technologies may be acceptable tools for providing social support to tuberculosis patients in rural Uganda: a parallel mixed-method study. Tuberc Res Treat 2020 Jan 07;2020:7401045-7401048 [FREE Full text] [doi: 10.1155/2020/7401045] [Medline: 31969998]
35. Yin J, Wang X, Zhou L, Wei X. The relationship between social support, treatment interruption and treatment outcome in patients with multidrug-resistant tuberculosis in China: a mixed-methods study. Trop Med Int Health 2018 Jun 16;23(6):668-677 [FREE Full text] [doi: 10.1111/tmi.13066] [Medline: 29691959]
36. Guzman K, Crowder R, Leddy A, Maraba N, Jennings L, Ahmed S, et al. Acceptability and feasibility of digital adherence technologies for drug-susceptible tuberculosis treatment supervision: a meta-analysis of implementation feedback. PLOS Digit Health 2023 Aug 15;2(8):e0000322 [FREE Full text] [doi: 10.1371/journal.pdig.0000322] [Medline: 37582066]
37. Mukora R, Ahumah B, Maraba N, Orrell C, Jennings L, Naidoo P, et al. Acceptability of using the medication monitor and experience of a differentiated care approach for TB treatment adherence among people living with TB in South Africa. PLOS Glob Public Health 2023 Oct 27;3(10):e0001885 [FREE Full text] [doi: 10.1371/journal.pgph.0001885] [Medline: 37889875]

Abbreviations

COREQ: Consolidated Criteria for Reporting Qualitative Research
DAT: digital adherence technology
MRRH: Mbarara Regional Referral Hospital
UTAUT: unified theory of acceptance and use of technology

XSL•FO
RenderX

Edited by A Kushniruk; submitted 07.04.23; peer-reviewed by B Noort, J Wang, B Thies; comments to author 12.08.23; revised version received 30.10.23; accepted 07.04.24; published 31.05.24.

Please cite as:
Musiimenta A, Tumuhimbise W, Atukunda E, Mugaba A, Linnemayr S, Haberer J
Digital Adherence Technologies Linked to Mobile Money Incentives for Medication Adherence Among People Living With Tuberculosis: Mixed Methods Feasibility and Acceptability Study
JMIR Hum Factors 2024;11:e47996
URL: https://humanfactors.jmir.org/2024/1/e47996
doi:10.2196/47996
PMID:38819905

©Angella Musiimenta, Wilson Tumuhimbise, Esther Atukunda, Aaron Mugaba, Sebastian Linnemayr, Jessica Haberer. Originally published in JMIR Human Factors (https://humanfactors.jmir.org), 31.05.2024. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on https://humanfactors.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

A Multicomponent Intervention (POSSIBLE) to Improve Perceived Risk for HIV Among Black Sexual Minority Men: Feasibility and Preliminary Effectiveness Pilot Study

Derek T Dangerfield II[1], PhD; Janeane N Anderson[2], MPH, PhD; Charleen Wylie[1]; Ricky Bluthenthal[3], PhD; Chris Beyrer[4], MD; Jason E Farley[5], RN, MPH, MSN, PhD

[1]Milken Institute School of Public Health, George Washington University, Washington, DC, United States
[2]University of Tennessee Health Science Center, Memphis, TN, United States
[3]Keck School of Medicine of USC, University of Southern California, Los Angeles, CA, United States
[4]Duke Global Health Institute, Durham, NC, United States
[5]Johns Hopkins School of Nursing, Baltimore, MD, United States

Corresponding Author:
Derek T Dangerfield II, PhD
Milken Institute School of Public Health
George Washington University
950 New Hampshire Avenue NW #308
Washington, DC, 20037
United States
Phone: 1 6673555188
Email: ddanger2@gwu.edu

Abstract

Background: Increased pre-exposure prophylaxis (PrEP) use is urgently needed to substantially decrease HIV incidence among Black sexual minority men. Low perceived risk for HIV (PRH) is a key unaddressed PrEP barrier for Black sexual minority men. Peers and smartphone apps are popular intervention tools to promote community health behaviors, but few studies have used these together in a multicomponent strategy. Therefore, we designed a multicomponent intervention called POSSIBLE that used an existing smartphone app called PrEPme (Emocha Mobile Health, Inc) and a peer change agent (PCA) to increase PRH as a gateway to PrEP.

Objective: This paper aims to describe the feasibility and preliminary impact of POSSIBLE on PRH and willingness to accept a PrEP referral among Black sexual minority men.

Methods: POSSIBLE was a theoretically guided, single-group, 2-session pilot study conducted among Black sexual minority men from Baltimore, Maryland between 2019 and 2021 (N=69). POSSIBLE integrated a PCA and the PrEPme app that allows users to self-monitor sexual risk behaviors and chat with the in-app community health worker to obtain PrEP service information. PRH was assessed using the 8-item PRH scale before and after baseline and follow-up study visits. At the end of each study visit, the PCA referred interested individuals to the community health worker to learn more about PrEP service options.

Results: The average age of participants was 32.5 (SD 8.1, range 19-62) years. In total, 55 (80%) participants were retained for follow-up at month 1. After baseline sessions, 29 (42%) participants were willing to be referred to PrEP services, 20 (69%) of those confirmed scheduled appointments with PrEP care teams. There were no statistically significant differences in PRH between baseline and follow-up visits (t_{122}=–1.36; P=.17).

Conclusions: We observed no statistically significant improvement in PRH between baseline and month 1. However, given the high retention rate and acceptability, POSSIBLE may be feasible to implement. Future research should test a statistically powered peer-based approach on PrEP initiation among Black sexual minority men.

Trial Registration: ClinicalTrials.gov NCT04533386; https://clinicaltrials.gov/study/NCT04533386

(JMIR Hum Factors 2024;11:e54739) doi:10.2196/54739

XSL•FO
RenderX

KEYWORDS

pre-exposure prophylaxis; PrEP; sexual health; peers; apps; community; mobile phone; HIV; sexual minority; minority communities; minority; Black; African American; peers; patient education; self-monitoring; treatment adherence; treatment participation; community; community health; mobile health; digital health; digital technology; digital interventions; smartphones

Introduction

Increased pre-exposure prophylaxis (PrEP) use is urgently needed to substantially decrease HIV incidence among Black sexual minority men who have an estimated 50% lifetime risk of HIV acquisition [1]. Data show racial inequities in uptake and adherence among sexual minority men who meet PrEP indications [2]. Socioecological factors such as side effect concerns, stigma, low health care access, and poor clinical experiences including patient-clinician communication remain substantial PrEP barriers for this community [3-5]. Low perceived risk for HIV (PRH) is also a key unaddressed PrEP barrier [6-9]. Some Black sexual minority men have low concerns about HIV acquisition because they think their current behaviors are lower risk than their past or their peers' behaviors and they do not associate HIV status with quality of life [6,10]. Other reasons for low PRH include being in a monogamous relationship and limited sexual activity [6,9,11]. However, low PRH inadequately reflects objective risks and disease severity for Black sexual minority men, given the high-community HIV incidence, suboptimal HIV care outcomes, and negative health consequences of long-term infection [7,8]. Therefore, multicomponent interventions that address PRH and other known barriers are needed to improve PrEP uptake.

Some HIV prevention interventions leverage in-group members as peer change agents (PCAs) to disseminate health-related information within the community for behavior change [12-15]. Peers are considered a valuable resource in marginalized communities to obtain health information, discuss taboo experiences, and help group members understand why behavior change could be beneficial [6,16,17]. Peers can facilitate behavior change because they have similar experiences, can address social barriers, and can improve health literacy [12,13,16,18]. PCAs are uniquely positioned to influence health behaviors because their roles as community members, patients, and health care paraprofessionals can build trust and lead others to credible information or clarify the information [17]. PCAs have improved behavioral health [12], HIV testing [19], medication adherence, and PrEP [13-15] for HIV prevention and could be effective interventionists among Black sexual minority men.

Other interventions have used smartphone apps as electronic diaries to reduce sexual risks through self-monitoring behaviors, which facilitates reflection [20-22]. Technology-based interventions could also be effective for Black sexual minority men because many of them use apps and other mobile devices for several purposes, including partner seeking, social network development, and health information [23,24]. Using apps for interventions could help Black sexual minority men circumvent the social and structural barriers to PrEP such as perceived judgment, stigma, and discrimination from clinicians. Since peers and smartphone apps are typically used independently in interventions, they could have a stronger impact if combined into a multicomponent health communication strategy because they could hypothetically reduce socioecological barriers for Black sexual minority men simultaneously.

Motivational interviewing (MI) is a communication approach in which a professional collaborates with individuals to activate their motivations to change behavior [17,25]. Some studies use "motivational interview consistent" interventions for HIV prevention because of their cost-effectiveness, brevity, and use of client interests for behavior change [26]. However, existing MI-based interventions to improve PRH or PrEP among Black sexual minority men are limited [18]. Additionally, factors known to drive PRH or PrEP use among Black sexual minority men by PCAs may not be fully leveraged in traditional MI-based interventions [27].

We designed a multicomponent intervention called POSSIBLE that used a PCA and an existing smartphone app called PrEPme (Emocha Mobile Health, Inc) [25] to increase PRH as an important cognitive gateway to PrEP uptake among Black sexual minority men. The intervention was guided by life course theory [28,29], the health belief model [30,31], and possible selves [6,32]. Life course theory suggests that timing, major life events, and age-related vulnerabilities impact sexual health behaviors [28,33]. The health belief model posits that risk perceptions catalyze behavior change [9,30]. Possible selves suggests that ideas of what individuals could or want to become can influence behavior [6,32,34]. Taken together, we hypothesized that a PCA who represented a "future self" could influence PRH at particular points along the life course of Black sexual minority men by cueing individuals to action through a review of their sexual risks in PrEPme and successfully encourage others to use PrEP having navigated similar social challenges [6].

This paper describes the feasibility and preliminary effects of POSSIBLE on PRH and subsequent willingness to accept a PrEP referral among Black sexual minority men. More information on the feasibility and effects of a multicomponent strategy using a PCA and smartphone app is needed to advance the promise of their combined effectiveness. Given the extreme racial disparity in PrEP use among sexual minority men [35,36], strategies that can increase PrEP uptake are still needed. Findings will provide insights into the usefulness of combining 2 popular HIV prevention interventions into a multicomponent strategy and elucidate the cognitive and cultural aspects of health decision-making among this vulnerable community.

Methods

Ethical Considerations

All study procedures were approved by the Johns Hopkins School of Medicine Institutional Review Board (IRB00241244). Oral informed consent was audio-recorded and documented in study folders prior to conducting baseline study visits. Participants were given study ID numbers, and identifiable

XSL•FO
RenderX

information (ie, names and data) was stored on private, password-protected servers. Participants were provided a US $50 electronic Amazon gift card for the baseline visit.

Study Design

POSSIBLE was a single-group, 2 session pilot intervention conducted between 2019 and 2021 that was refined using the ADAPT-ITT (Assessment, Decision, Administration, Production, Topical Experts, Integration, Training, and Testing) model [37]. Formative research was conducted to refine key aspects of the intervention approach such as the usefulness of the app-based diary and PCA characteristics among Black sexual minority men of different age groups [6].

Peer Change Agent

POSSIBLE incorporated a PCA who was matched with participants' key demographic and cultural characteristics as guided by theory, previous studies, and formative research identifying preferences among Black sexual minority men [6,38]. Specifically, studies suggested that the PCA should have similar experiences as Black sexual minority men (eg, navigating interactions with romantic or sexual partners, clinicians, and health insurance) and be a "future self" to whom they could aspire [6,38]. Therefore, the principal investigator (PI to DTD), used Descovy for PrEP and served the intervention as the PCA. Further details regarding the experience and dual role of the PI serving as a PrEP-using PCA have been published in an autoethnography [17].

PrEPme Smartphone App

PrEPme was designed for Maryland users to obtain statewide PrEP service information and navigation support from an app-based community health worker (CHW) [25]. PrEPme also allows users to self-monitor sexual risk behaviors, view a graph of sexual risk behaviors by week and month, and chat with the CHW in the app to obtain PrEP service information [6].

Linkage to PrEP Care

A CHW supervised by a nurse case manager within the Center for Infectious Disease and Nursing Innovation at the Johns Hopkins School of Nursing (previously known as the REACH Initiative) provided navigation services including reviewing eligibility for or access to medical insurance, identifying preferred clinic locations, and arranging appointment and scheduling activities. Occasionally, the PCA referred participants or helped schedule appointments at PrEP service organizations for individuals who wanted to avoid interfacing with additional staff associated with a medical research institution.

Study Procedures

Study Enrollment

A research assistant screened individuals who were interested in the study for eligibility via phone. Eligible individuals were emailed an informed consent form, details regarding their scheduled web-based baseline appointment, and an electronic survey assessing demographic and behavioral characteristics and PRH. Individuals were given the opportunity to ask the research team (including the PI) questions regarding the study via phone or email prior to their baseline visit.

Web-Based Baseline Study Visit

Due to COVID-19, baseline and follow-up study visits were conducted via Zoom (Zoom Video Communications). Prior to the baseline visit, the PI or PCA provided participants an additional opportunity to ask questions regarding the study and obtained oral informed consent via Zoom [39]. The script guided the PCA to obtain information regarding participants' lifestyles, personal goals and values, relative HIV risks, and PRH, then tailor health communication based on their responses to influence PRH and encourage PrEP use regardless of participants' reported behaviors [16,26,27,29]. Example questions included, "How would being diagnosed with HIV impact your goals?" and "Given that research suggests that 50% of Black sexual minority men will get HIV despite the fact that they use condoms more than other people, how likely do you think you will get it?" [6,40]. The script also provided opportunities for the PCA to address HIV or PrEP misinformation and disclose PrEP use to share experiences managing potential side effects, challenges disclosing use to romantic or sexual partners, and empathy regarding participants' stigmatizing clinical experiences [17,40].

At the end of the session, the PCA referred interested individuals to the CHW as described earlier. Individuals who declined referrals to the CHW were provided alternative service locations for PrEP care and referred upon request. Baseline visits lasted between 45 and 60 minutes (accounting for informed consent, rapport building, 15- to 20-minute conversation, and PrEP navigation for those who were interested), at the end of which participants were asked to download PrEPme to record their sexual risk behaviors in its app-based diary for 1 month. Baseline study visit procedures and effects have been published [40].

Web-Based Follow-Up Visit

In the second session, the PCA reviewed the PrEPme diary with participants and conducted another MI-consistent conversation to explore relative HIV risk behaviors, review behavioral alignment with goals and values, and reassess their PrEP interests. At the end of the session, the PCA referred interested individuals to the CHW as described earlier. Individuals who declined referrals to the CHW were provided alternative service locations for PrEP care and referred upon request. Follow-up visits lasted between 20 and 30 minutes, and participants were provided another US $50 electronic Amazon gift card for completing follow-up visits regardless of reported app use.

PrEP Referral

All participants were first referred a CHW at the Johns Hopkins School of Nursing who could help navigate them to PrEP services. Participants who were interested in case management from the CHW were linked to services of their choice. Individuals who declined referrals to the CHW were offered direct referrals by the PCA who reached out to the requested case management services to help schedule appointments.

XSL•FO
RenderX

Satisfaction Surveys

Participants completed a satisfaction survey that also assessed their PRH at the end of the baseline session prior to downloading PrEPme, then again at the end of their follow-up appointment.

Participants

Participants were recruited using a combination of active and passive strategies [39,41] in Baltimore, Maryland, and were eligible based upon the following self-reported criteria: Black or African American race, identifying as a cisgender person, being 18 years and older of age, same-sex attraction to men, HIV-negative, and having oral or anal sex with ≥1 male partner in the previous 6 months.

Measures

This concept was assessed using the 8-item PRH scale from Napper et al [42] before and after the baseline visit. Sample questions included items assessing concerns about HIV and perceived likelihoods of infection. Total possible scores ranged from 10 to 40, higher scores indicate greater PRH. The scale was found to be reliable (8 items, α=.78).

Data Analysis

Paired 1-tailed *t* tests were used to examine changes in PRH after the end of the study. Descriptive statistics were used to explore the proportion of participants who were referred to services and scheduled a PrEP appointment after baseline.

Results

A total of 291 individuals were screened for the study, 93 of whom were eligible. Among eligible individuals, 69 participated and 55 (80%) were retained for follow-up at month 1. Table 1 describes the sociodemographic characteristics and PrEP referral willingness among participants. The average age of participants was 32.5 (SD 8.1, range 19-62) years. Additionally, 52 (75%) identified as gay, 11 (16%) identified as bisexual, 51 (74%) reported being employed full-time or part-time at baseline, 58 (84%) reported having insurance coverage, 54 (78%) reported being single, and 32 (47%) reported ever having a sexually transmitted infection. After baseline sessions, 29 (42%) participants were willing to be referred to PrEP services, and 20 (69%) of them confirmed scheduled appointments with PrEP care teams.

Regarding the use of the mobile app–based diary, 17 (31%) follow-up participants reported recording an entry every week prior to their follow-up appointment, 3 (5%) reported using the app for half of the weeks, and 6 (11%) reported that they did not use the app at all. In total, 11 (20%) reported initiating PrEP prior to follow-up, and 15 (27%) of follow-up participants were willing to be referred to PrEP services. There were no statistically significant differences in mean PRH scores between baseline (21.2, SD 5.5) and follow-up (23.6, SD 5.7) visits (t_{122}=–1.36; *P*=.17).

XSL•FO
RenderX

Table 1. Sociodemographic characteristics and PrEP[a] referral willingness among Black sexual minority men in POSSIBLE 2019-2021 (N=69).

	Baseline	Month 1 follow-up
Age (years)		
Mean (SD)	32.5 (8.1)	32.7 (7.7)
Range	19-62	19-50
<35 years, n (%)	49 (71)	40 (73)
Sexual orientation, n (%)		
Homosexual, gay, same gender-loving	52 (75)	42 (76)
Bisexual	11 (16)	7 (13)
Other	6 (9)	6 (11)
Employment status, n (%)		
Full-time	44 (64)	36 (65)
Part-time	7 (10)	6 (11)
Unemployed	13 (19)	11 (20)
Other	5 (7)	2 (4)
Highest level of education, n (%)		
Grade 11 or less	5 (7)	1 (2)
Grade 12 or GED[b]	10 (14)	11 (20)
Associate degree	2 (3)	0 (0)
Some college	10 (14)	7 (13)
Bachelor degree	24 (35)	16 (29)
More than bachelor degree	18 (26)	20 (36)
Health care coverage, n (%)	58 (84)	51 (93)
Annual gross income (US $), n (%)		
Less than $20,000	15 (22)	15 (27)
Between $30,000 and $40,000	8 (11)	12 (22)
Between $40,000 and $50,000	9 (13)	7 (13)
Between $50,000 and $60,000	6 (9)	4 (7)
More than $60,000	23 (33)	17 (31)
Relationship status, n (%)		
Single or not in a relationship	54 (78)	45 (82)
In a committed relationship	9 (13)	7 (13)
Married	2 (3)	0 (0)
Other	2 (3)	3 (5)
STI[c] history past 6 months, n (%)	32 (47)	42 (76)
Drug use before sex past 6 months, n (%)	47 (68)	39 (70)
Willingness to be referred to PrEP services, n (%)	29 (42)	15 (27)
PrEP appointment scheduled	20 (69)	__[d]
Initiated PrEP prior to month 1	—	11 (20)

[a]PrEP: pre-exposure prophylaxis.
[b]GED: General Educational Diploma.
[c]STI: sexually transmitted infection.
[d]Not available.

XSL•FO
RenderX

Discussion

Principal Findings

This study explored the feasibility and preliminary effectiveness of POSSIBLE, a multicomponent intervention using a PCA and mobile app–based diary to improve PRH among Black sexual minority men. Given the high retention rate, POSSIBLE may be a feasible multicomponent strategy to implement among Black sexual minority men. We found improvements in PRH after baseline sessions [40]. However, we observed no statistically significant improvement in PRH after intervention from baseline to month 1.

We observed relatively low PRH scores at baseline and month 1 follow-up. Analyses showed that the PCA increased baseline PRH scores [40]. The effects of the intervention may have been maximized in the baseline session such that the addition of the app for reflexivity could not increase scores from baseline to month 1. Other studies have found that competing survival priorities supersede HIV-related concerns in the lives of Black sexual minority men. The shift from HIV as a "death sentence" in the early days of the epidemic to its positioning as a manageable chronic health condition could be a key reason for low PRH and for why perceived risk did not change from baseline to month 1. Black sexual minority men may also consider their current behaviors relatively safer than their past or their peers' behaviors as found in previous studies [6,10]. Some may appraise their vulnerability based upon their most recent behaviors, which may not have involved condomless sex or drug use in the month of the intervention, which was conducted partly during the height of the global COVID-19 pandemic.

We also observed relatively low use of the app-based diary in PrEPme. Mobile health interventions have been successful largely because of the convenience of the intervention within smartphone apps. However, PrEPme seemingly did not add value to the PCA intervention component. Studies suggest that aspects such as aesthetics, social networking ability, and gamification impact Black sexual minority men's use of app-based HIV prevention apps [24,43]. The self-monitoring feature of the app-based diary could be refined for gamification and cultural responsiveness. PrEPme also did not maximize the power of health communication to tailor messaging. However, we are unable to identify reasons for nonuse. Qualitative insights or participant feedback could help in identifying barriers to app use. In light of previous analyses showing baseline effects on PRH [40], adding the mobile app–based diary may not be necessary in the presence of an effective PCA.

The intervention did observe relatively high proportions of willingness to accept PrEP referral and initiation, which could be attributed to the interpersonal dynamics between participants and the PCA [17]. Studies consistently show that peers and social networks are important, trusted sources of health information and effective interventionists among marginalized communities, including Black sexual minority men [16,38,44,45]. Studies also show that Black sexual minority men are more willing to initiate PrEP if their peers are using it. The usefulness of communicating with a PCA may not have been outweighed by the convenience of technology. PRH may not necessarily be the primary motivation for PrEP referral willingness or initiation among Black sexual minority men.

Despite the feasibility of this intervention, using a PCA to catalyze PRH and PrEP initiation among Black sexual minority men is not without challenges. DTD described internal conflicts regarding honoring participants' disinterest in PrEP versus professional goals to increase uptake for HIV prevention in an autoethnography [17]. Additionally, managing discussions regarding side effects with Black sexual minority men whose health histories the PI or PCA was unfamiliar with or unqualified to discuss is important to consider in this peer-based approach. Concerns that being a PCA could overshadow professionalism as a researcher and health care professional were also salient [17]. PCAs should be trained to manage insider-outsider dynamics as an interventionist among Black sexual minority men, engage in active listening, and communicate with care [17,46]. Some qualities may not be able to be provided in training such as the shared social experiences and vulnerabilities of being Black sexual minority men.

Limitations

Study limitations include insufficient sample size to detect effect sizes. Causal inferences cannot be drawn, and effectiveness cannot be established with a pre-post single-group design. It is also possible that unstudied external factors could have produced the changes observed. All data were self-reported. A larger trial is needed to definitively establish the effects of the intervention, including biological confirmation of PrEP use beyond self-report.

Conclusions

Future research should test a statistically powered peer-based approach on PrEP initiation among Black sexual minority men. Psychometric tests should also be conducted to identify culturally relevant concepts of HIV risk and PrEP motivation for Black sexual minority men. Qualitative research should also clarify how app-based sexual risk diaries have unintended consequences of triggering self-stigma and shame versus informed decision-making [6]. Targeted studies among young Black sexual minority men younger than 35 years of age should be conducted, given their high HIV incidence and low PrEP uptake. Studies show age cohort differences regarding the needs and vulnerabilities among Black sexual minority men such that peers may be a more effective behavior change mechanism for younger men [33,47,48]. If effectively implemented, the person-centered approach of a PrEP-using PCA approach could lead to substantial community-level impact for Black sexual minority men because their needs are not the same.

XSL•FO
RenderX

Acknowledgments

This research was supported by the National Institute of Mental Health Grant (#K01MH11849; principal investigator to DTD). This research has been facilitated by the infrastructure and resources provided by the Johns Hopkins University Center for AIDS Research, a National Institutes of Health (NIH) funded program (1P30AI094189), which is supported by the following NIH cofunding and participating institutes and centers: the National Institute of Allergy and Infectious Disease, the National Cancer Institute, the National Institute of Child Health and Human Development, the National Health, Lung, and Blood Institute, the National Institute on Drug Abuse, the National Institute on Aging, the National Institute of General Medical Sciences, the National Institute of Diabetes and Digestive Kidney Diseases, and the National Institute of Minority Health and Health Disparities. The content is solely the responsibility of the authors and does not necessarily represent the official views of the NIH. PrEPme was developed with funding from several sources including The Johns Hopkins School of Nursing REACH Initiative, the Gilead Science Unrestricted Educational Grant, and the Baltimore City Health Department.

Conflicts of Interest

JEF holds the technology transfer license with Johns Hopkins University for PrEPme. The app was developed in collaboration with Emocha Mobile Health, Inc.

References

1. Hess KL, Crepaz N, Rose C, Purcell D, Paz-Bailey G. Trends in sexual behavior among men who have sex with men (MSM) in high-income countries, 1990-2013: a systematic review. AIDS Behav 2017;21(10):2811-2834 [FREE Full text] [doi: 10.1007/s10461-017-1799-1] [Medline: 28555317]
2. Schumacher C, Tao X, Chandran A, Fields E, Price A, Greenbaum A, IMPACT Partner Collaborative. Reaching those most at risk for HIV acquisition: evaluating racial/ethnic disparities in the preexposure prophylaxis care continuum in Baltimore City, Maryland. J Acquir Immune Defic Syndr 2021;87(5):1145-1153 [FREE Full text] [doi: 10.1097/QAI.0000000000002712] [Medline: 33883471]
3. Quinn K, Dickson-Gomez J, Zarwell M, Pearson B, Lewis M. "A Gay Man and a Doctor are Just like, a Recipe for Destruction": how racism and homonegativity in healthcare settings influence PrEP uptake among young Black MSM. AIDS Behav 2019;23(7):1951-1963 [FREE Full text] [doi: 10.1007/s10461-018-2375-z] [Medline: 30565092]
4. Cahill S, Taylor SW, Elsesser SA, Mena L, Hickson D, Mayer KH. Stigma, medical mistrust, and perceived racism may affect PrEP awareness and uptake in black compared to white gay and bisexual men in Jackson, Mississippi and Boston, Massachusetts. AIDS Care 2017;29(11):1351-1358 [FREE Full text] [doi: 10.1080/09540121.2017.1300633] [Medline: 28286983]
5. Babel RA, Wang P, Alessi EJ, Raymond HF, Wei C. Stigma, HIV risk, and access to HIV prevention and treatment services among men who have sex with men (MSM) in the United States: a scoping review. AIDS Behav 2021;25(11):3574-3604 [FREE Full text] [doi: 10.1007/s10461-021-03262-4] [Medline: 33866444]
6. Dangerfield II DT, Anderson JN, Wylie C, Arrington-Sanders R, Bluthenthal RN, Beyrer C, et al. Refining a multicomponent intervention to increase perceived HIV risk and PrEP initiation: focus group study among Black sexual minority men. JMIR Form Res 2022;6(8):e34181 [FREE Full text] [doi: 10.2196/34181] [Medline: 35947442]
7. Blumenthal J, Jain S, Mulvihill E, Sun S, Hanashiro M, Ellorin E, et al. Perceived versus calculated HIV risk: implications for pre-exposure prophylaxis uptake in a randomized trial of men who have sex with men. J Acquir Immune Defic Syndr 2019;80(2):e23-e29 [FREE Full text] [doi: 10.1097/QAI.0000000000001888] [Medline: 30422909]
8. Oostrom L, Rosentel K, Motley D, Hill BJ. Discordance in objective and self-perceived HIV risk: a potential barrier to pre-exposure prophylaxis in young gay and bisexual men. J Assoc Nurses AIDS Care 2020;31(1):103-109 [FREE Full text] [doi: 10.1097/JNC.0000000000000137] [Medline: 31613824]
9. Felsher M, Szep Z, Krakower D, Martinez-Donate A, Tran N, Roth AM. "I Don't Need PrEP Right Now": a qualitative exploration of the barriers to PrEP care engagement through the application of the health belief model. AIDS Educ Prev 2018;30(5):369-381 [FREE Full text] [doi: 10.1521/aeap.2018.30.5.369] [Medline: 30332306]
10. Gallagher T, Link L, Ramos M, Bottger E, Aberg J, Daskalakis D. Self-perception of HIV risk and candidacy for pre-exposure prophylaxis among men who have sex with men testing for HIV at commercial sex venues in New York City. LGBT Health 2014;1(3):218-224. [doi: 10.1089/lgbt.2013.0046] [Medline: 26789715]
11. Huang W, Lockard A, Kelley CF, Serota DP, Rolle CM, Sullivan PS, et al. From declining PrEP to PrEP initiation as "first nature"—what changes PrEP initiation decisions among young, Black MSM. AIDS Care 2022;34(3):284-293 [FREE Full text] [doi: 10.1080/09540121.2021.1960946] [Medline: 34369230]
12. Weeks MR, Convey M, Dickson-Gomez J, Li J, Radda K, Martinez M, et al. Changing drug users' risk environments: peer health advocates as multi-level community change agents. Am J Community Psychol 2009;43(3-4):330-344 [FREE Full text] [doi: 10.1007/s10464-009-9234-z] [Medline: 19326208]
13. Young LE, Schumm P, Alon L, Bouris A, Ferreira M, Hill B, et al. PrEP Chicago: a randomized controlled peer change agent intervention to promote the adoption of pre-exposure prophylaxis for HIV prevention among young Black men who have sex with men. Clin Trials 2018;15(1):44-52 [FREE Full text] [doi: 10.1177/1740774517730012] [Medline: 28862483]

XSL•FO
RenderX

14. Turpin RE, Hawthorne DJ, Rosario AD. Pre-exposure prophylaxis interventions among Black sexual minority men: a systematic literature review. Int J Environ Res Public Health 2022;19(4):1934 [FREE Full text] [doi: 10.3390/ijerph19041934] [Medline: 35206120]
15. Harawa NT, Schrode KM, McWells C, Weiss RE, Hilliard CL, Bluthenthal RN. Small randomized controlled trial of the new passport to wellness HIV prevention intervention for Black men who have sex With men (BMSM). AIDS Educ Prev 2020;32(4):311-324 [FREE Full text] [doi: 10.1521/aeap.2020.32.4.311] [Medline: 32897130]
16. Quinn KG, Christenson E, Spector A, Amirkhanian Y, Kelly JA. The influence of peers on PrEP perceptions and use among young Black gay, bisexual, and other men who have sex with men: a qualitative examination. Arch Sex Behav 2020;49(6):2129-2143 [FREE Full text] [doi: 10.1007/s10508-019-01593-x] [Medline: 32016815]
17. Dangerfield II DT. A principal investigator as a PrEP-using peer change agent for HIV prevention among Black gay and bisexual men: an autoethnography. Int J Environ Res Public Health 2023;20(7):5309 [FREE Full text] [doi: 10.3390/ijerph20075309] [Medline: 37047925]
18. Dennis C. Peer support within a health care context: a concept analysis. Int J Nurs Stud 2003;40(3):321-332. [doi: 10.1016/s0020-7489(02)00092-5] [Medline: 12605954]
19. Fuqua V, Chen Y, Packer T, Dowling T, Ick TO, Nguyen B, et al. Using social networks to reach Black MSM for HIV testing and linkage to care. AIDS Behav 2012;16(2):256-265. [doi: 10.1007/s10461-011-9918-x] [Medline: 21390535]
20. Hensel DJ, Fortenberry JD, Harezlak J, Craig D. The feasibility of cell phone based electronic diaries for STI/HIV research. BMC Med Res Methodol 2012;12:75 [FREE Full text] [doi: 10.1186/1471-2288-12-75] [Medline: 22691189]
21. Glick SN, Winer RL, Golden MR. Web-based sex diaries and young adult men who have sex with men: assessing feasibility, reactivity, and data agreement. Arch Sex Behav 2013;42(7):1327-1335 [FREE Full text] [doi: 10.1007/s10508-012-9984-9] [Medline: 22926939]
22. Stalgaitis C, Glick SN. The use of web-based diaries in sexual risk behaviour research: a systematic review. Sex Transm Infect 2014;90(5):374-381. [doi: 10.1136/sextrans-2013-051472] [Medline: 24723619]
23. Fields EL, Morgan A, Sanders RA, Jennings J. “I Met Him On Jack’D”—exploring HIV risk perceptions and risk mitigation strategies of young Black MSM meeting sex partners on geosocial networking apps. J Adolesc Health 2017;60(2):S8. [doi: 10.1016/j.jadohealth.2016.10.037]
24. Holloway IW, Winder TJ, Lea III CH, Tan D, Boyd D, Novak D. Technology use and preferences for mobile phone-based HIV prevention and treatment among Black young men who have sex with men: exploratory research. JMIR Mhealth Uhealth 2017;5(4):e46 [FREE Full text] [doi: 10.2196/mhealth.6436] [Medline: 28408360]
25. Farley JE, Dangerfield II DT, LaRicci J, Sacamano P, Heidari O, Lowensen K, et al. Community engagement and linkage to care efforts by peer community-health workers to increase PrEP uptake among sexual minority men. Public Health Nurs 2021;38(5):818-824 [FREE Full text] [doi: 10.1111/phn.12887] [Medline: 33749022]
26. Rollnick S, Miller WR. What is motivational interviewing? Behav Cogn Psychother 1995;23(4):325-334. [doi: 10.1017/S135246580001643X]
27. Dangerfield II DT, Davis G, Pandian V, Anderson JN. Using motivational interviewing to increase HIV PrEP initiation and adherence: a scoping review. Prev Sci 2023;24(7):1365-1375 [FREE Full text] [doi: 10.1007/s11121-023-01554-w] [Medline: 37249729]
28. Elder GH. The life course as developmental theory. Child Dev 1998;69(1):1-12. [Medline: 9499552]
29. Elder GH. Human lives in changing societies: life course and developmental insights. In: Cairns RB, Elder GH, Costello EJ, editors. Developmental Science. North Carolina: Cambridge University Press; 1996:31-62.
30. Rosenstock IM. The health belief model and preventive health behavior. Health Educ Monogr 1974;2(4):354-386. [doi: 10.1177/109019817400200405]
31. Rosenstock IM. Historical origins of the health belief model. Health Educ Monogr 1974;2(4):328-335. [doi: 10.1177/109019817400200403]
32. Markus H, Nurius P. Possible selves. Am Psychol 1986;41(9):954-969. [doi: 10.1037/0003-066X.41.9.954]
33. Dangerfield II DT, Smith LR, Anderson JN, Bruce OJ, Farley J, Bluthenthal R. Sexual positioning practices and sexual risk among Black gay and bisexual men: a life course perspective. AIDS Behav 2018;22(6):1919-1931 [FREE Full text] [doi: 10.1007/s10461-017-1948-6] [Medline: 29079948]
34. Bak W. Possible selves: implications for psychotherapy. Int J Ment Health Addict 2015;13(5):650-658 [FREE Full text] [doi: 10.1007/s11469-015-9553-2] [Medline: 26366132]
35. Chu BA, Castellanos ER, Gonzales MM, Gaither TW. Social determinants of self-reported pre-exposure prophylaxis use among a national sample of US men who have sex with men. Sex Transm Infect 2020;96(5):368-374. [doi: 10.1136/sextrans-2019-054177] [Medline: 31601642]
36. Kanny D, Jeffries WL, Chapin-Bardales J, Denning P, Cha S, Finlayson T, et al. Racial/ethnic disparities in HIV preexposure prophylaxis among men who have sex with men—23 urban areas, 2017. MMWR Morb Mortal Wkly Rep 2019;68(37):801-806 [FREE Full text] [doi: 10.15585/mmwr.mm6837a2] [Medline: 31536484]
37. Wingood GM, DiClemente RJ. The ADAPT-ITT model: a novel method of adapting evidence-based HIV interventions. J Acquir Immune Defic Syndr 2008 Mar 01;47(Suppl 1):S40-S46. [doi: 10.1097/QAI.0b013e3181605df1] [Medline: 18301133]

XSL•FO
RenderX

38. Dangerfield II DT, Harawa NT, McWells C, Hilliard C, Bluthenthal RN. Exploring the preferences of a culturally congruent, peer-based HIV prevention intervention for black men who have sex with men. Sex Health 2018;15(5):424-430 [FREE Full text] [doi: 10.1071/SH18057] [Medline: 30185352]
39. Dangerfield II DT, Wylie C, Anderson JN. Conducting virtual, synchronous focus groups among Black sexual minority men: qualitative study. JMIR Public Health Surveill 2021;7(2):e22980 [FREE Full text] [doi: 10.2196/22980] [Medline: 33427671]
40. Dangerfield II DT, Anderson JN. A scripted, PrEP-using peer change agent improves perceived risk for HIV and willingness to accept referrals quickly among Black sexual minority men: preliminary findings from POSSIBLE. AIDS Behav 2024. [doi: 10.1007/s10461-024-04313-2] [Medline: 38551719]
41. White JJ, Dangerfield II DT, Grieb SM. Methodological considerations for conducting focus groups in HIV prevention research among Black men who have sex with men. Public Health Nurs 2019;36(3):439-445 [FREE Full text] [doi: 10.1111/phn.12592] [Medline: 30761620]
42. Napper LE, Fisher DG, Reynolds GL. Development of the perceived risk of HIV scale. AIDS Behav 2012;16(4):1075-1083 [FREE Full text] [doi: 10.1007/s10461-011-0003-2] [Medline: 21785873]
43. Goldenberg T, McDougal SJ, Sullivan PS, Stekler JD, Stephenson R. Preferences for a mobile HIV prevention app for men who have sex with men. JMIR Mhealth Uhealth 2014 Oct 29;2(4):e47 [FREE Full text] [doi: 10.2196/mhealth.3745] [Medline: 25355249]
44. Kelly JA. Popular opinion leaders and HIV prevention peer education: resolving discrepant findings, and implications for the development of effective community programmes. AIDS Care 2004;16(2):139-150. [doi: 10.1080/09540120410001640986] [Medline: 14676020]
45. Kelly JA, St Lawrence JS, Diaz YE, Stevenson LY, Hauth AC, Brasfield TL, et al. HIV risk behavior reduction following intervention with key opinion leaders of population: an experimental analysis. Am J Public Health 1991;81(2):168-171. [doi: 10.2105/ajph.81.2.168] [Medline: 1990853]
46. Dangerfield II DT, Cooper J, Heidari O, Allen S, Winder TJ, Lucas GM. Nursing and health care preferences among opioid and stimulant using Black sexual minority men: an exploratory study. J Assoc Nurses AIDS Care 2021;32(5):e30-e39. [doi: 10.1097/JNC.0000000000000201] [Medline: 32809988]
47. Dangerfield II DT, Heidari O, Cooper J, Allen S, Lucas GM. Motivations for opioid and stimulant use among drug using black sexual minority men: a life course perspective. Drug Alcohol Depend 2020;215:108224 [FREE Full text] [doi: 10.1016/j.drugalcdep.2020.108224] [Medline: 32777690]
48. Dangerfield II DT, Harawa NT, Fernandez MI, Hosek S, Lauby J, Joseph H, et al. Age cohort differences in sexual behaviors among Black men who have sex with men and women. J Sex Res 2018;55(8):1012-1021 [FREE Full text] [doi: 10.1080/00224499.2017.1423016] [Medline: 29377734]

Abbreviations

ADAPT-ITT: Assessment, Decision, Administration, Production, Topical Experts, Integration, Training, and Testing
CHW: community health worker
MI: motivational interviewing
PCA: peer change agent
PI: principal investigator
PrEP: pre-exposure prophylaxis
PRH: perceived risk for HIV

Edited by C Jacob; submitted 20.11.23; peer-reviewed by S Harrison, G Sebastian, S Markham; comments to author 31.01.24; revised version received 17.02.24; accepted 22.04.24; published 11.06.24.

Please cite as:
Dangerfield II DT, Anderson JN, Wylie C, Bluthenthal R, Beyrer C, Farley JE
A Multicomponent Intervention (POSSIBLE) to Improve Perceived Risk for HIV Among Black Sexual Minority Men: Feasibility and Preliminary Effectiveness Pilot Study
JMIR Hum Factors 2024;11:e54739
URL: https://humanfactors.jmir.org/2024/1/e54739
doi:10.2196/54739
PMID:38861707

XSL•FO
RenderX

©Derek T Dangerfield II, Janeane N Anderson, Charleen Wylie, Ricky Bluthenthal, Chris Beyrer, Jason E Farley. Originally published in JMIR Human Factors (https://humanfactors.jmir.org), 11.06.2024. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on https://humanfactors.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Characteristics, Opportunities, and Challenges of Osteopathy Based on the Perceptions of Osteopaths in Austria: Qualitative Interview Study

Jonas Manschel[1], MSc; Jan Porthun[2], MMMSc; Ulrike Winkler[3], Dr phil; Jean Marie A T Beuckels[1,4], Dr rer medic; David Martin[1,5], Dr med

[1]Institute of Integrative Medicine, Department of Medicine, Health Faculty, University Witten/Herdecke, Herdecke, Germany
[2]Norwegian University of Science and Technology, Campus Gjøvik, Gjøvik, Norway
[3]Bundeswehr University Munich, Neubiberg, Germany
[4]Department of Osteopathy, Faculty of Health and Social Sciences, Hochschule Fresenius, Munich, Germany
[5]Tübingen University Children's Hospital, Tübingen, Germany

Corresponding Author:
Jan Porthun, MMMSc
Norwegian University of Science and Technology, Campus Gjøvik
Teknologivn 22
Gjøvik, 2815
Norway
Phone: 47 95007916
Email: jan.porthun@ntnu.no

Abstract

Background: There are no uniform regulations for the osteopathic profession in Europe. It is subject to country-specific regulations defining who shall be allowed to practice osteopathy and which qualification shall be required. In recent years, legal regulations have been established in several European countries for the profession of osteopathy; however, these are also still pending for Austria. Currently, physiotherapists and physicians with osteopathic training are practicing osteopathy in Austria.

Objective: This study aims to examine the characteristics, challenges, and opportunities of osteopaths in Austria.

Methods: Guideline-based interviews with osteopaths (N=10) were conducted. The different research questions were examined using a qualitative content analysis.

Results: The study provided a differentiated insight into the professional situation of osteopaths in Austria. The most important result was that all interviewees unanimously supported a legal regulation of their profession. However, owing to their different professional self-image—on the one hand, individuals working on a structural basis, and, on the other hand, individuals working on a cranial or biodynamic basis—they were able to imagine a uniform professional regulation only to a limited extent. Additional topics for the interviewed osteopaths in Austria were the quality assurance of training and the urgent need for scientific research. Furthermore, the study also dealt with the influence of the COVID-19 pandemic on daily practice and on education and training in osteopathy.

Conclusions: This study is a pioneering study with regard to systematic basic research on osteopathy in Austria. The obtained results and the newly acquired research questions not only have the potential to serve as a basis for further studies but also provide insight into the working and professional situation of osteopaths in Austria for universities, schools, professional associations, politics, and—last but not least—all interested parties.

International Registered Report Identifier (IRRID): RR2-10.2196/15399

(JMIR Hum Factors 2024;11:e45302) doi:10.2196/45302

KEYWORDS

osteopathy; osteopath; osteopaths; osteopathic profession; health care system; Austria

XSL•FO
RenderX

Introduction

Osteopathy is a manual treatment method, the principles of which are based on its own philosophy and the consideration and treatment of special structure-function relationships in the human body [1]. Since osteopathy was established and as long as it has been applied, both its methods and the professional competence of osteopaths have been the subject of controversy among medical and therapeutic specialists. The European Committee for Standardization has defined osteopathy as a holistic, patient-centered, manual treatment method based on the interactions between the structure and function of the body and the body's self-healing ability [2].

There is no uniform European or international regulation regarding who is allowed to practice osteopathy and which qualifications are required. However, an increasing number of European countries are developing occupational laws for osteopaths. So far, 12 European countries including Cyprus, Denmark, Finland, France, Iceland, Liechtenstein, Luxembourg, Malta, Norway, Portugal, Switzerland, and the United Kingdom have adopted legal regulations regarding the practice of osteopaths [3].

In the German-speaking countries, there is no uniform picture of the profession. In contrast to Switzerland, a legal basis for the profession of osteopathy does not exist in Germany or Austria. In Austria, physicians and physiotherapists trained in osteopathy practice as osteopaths. Physicians are allowed to practice osteopathy without any restrictions, whereas physiotherapists are only allowed to practice osteopathy upon medical assignment [4]. According to the Austrian Society for Osteopathy (Österreichische Gesellschaft für Osteopathie; OEGO), approximately 2000 osteopaths practiced in Austria in 2022.

Studies about osteopathic identity are progressing internationally. However, the various legal regulations and intraprofessional conflicts make it difficult to perceive a collective identity [5]. Especially in countries where osteopathy is not regulated by law, the data about osteopathic practitioners are considered to be weak. However, quantitative studies that have surveyed the population of osteopaths with regard to work status, training, professional identity, or characteristics of clinical practice such as the typical patient profile and the use of diagnostic and treatment modalities exist already. In Austria, 2 surveys of osteopaths have been conducted in the past as part of final theses [6,7]. In 2022, the results of the Osteopathic Practitioners, Estimates, and Rates survey were also published for Austria, thus creating a solid data basis about osteopathic practitioners in Austria for the first time. The typical osteopath was defined as female, aged between 40 and 49 years, self-employed, worked before as a physiotherapist, trained in osteopathy part time, and successfully completed a master's degree [8]. However, there is a lack of studies with qualitative designs to capture and examine the work of osteopaths in German-speaking countries in more detail.

The overall aim of this study was to make a substantial contribution to the largely unexplored profession of osteopathy in the German-speaking countries. Structured, basic research was necessary to implement this project. The first steps were taken in the framework of the study, "Characteristics, Opportunities, and Challenges of Osteopathy (COCO) in the Perceptions of Osteopaths in Germany, Austria, and Switzerland: Protocol for a Comprehensive Mixed Methods Study." The study protocol was published in *JMIR Research Protocols* in 2019. The Characteristics, Opportunities, and Challenges of Osteopathy (COCO) project investigates how osteopaths in Germany, Austria, and Switzerland distinguish themselves from other medical professions and the characteristics of their work.

This study is a partial study of the COCO project, with a focus on the situation of osteopaths in Austria. Osteopaths practicing in Austria were asked about their professional profile and their professional practice. The following questions were of particular interest: (1) How do osteopaths from Austria describe osteopathy? (2) What are the challenges faced by osteopaths in Austria? and (3) What opportunities do the interviewees see for osteopathy in Austria?

Methods

Design

This qualitative study included the planning and implementation of guideline-based interviews with osteopaths practicing in Austria. Subsequently, a qualitative content analysis was performed according to Mayring [9]. A qualitative research design was selected to obtain questions relevant to the project that had not been considered before and views about the topic that had not yet been taken into consideration. The target of this qualitative partial study was the development of hypotheses. Accordingly, a relatively small sample of 8 to 10 participants could be used, because the results obtained shall be examined in subsequent studies with respect to their general validity using a quantitative study design [10].

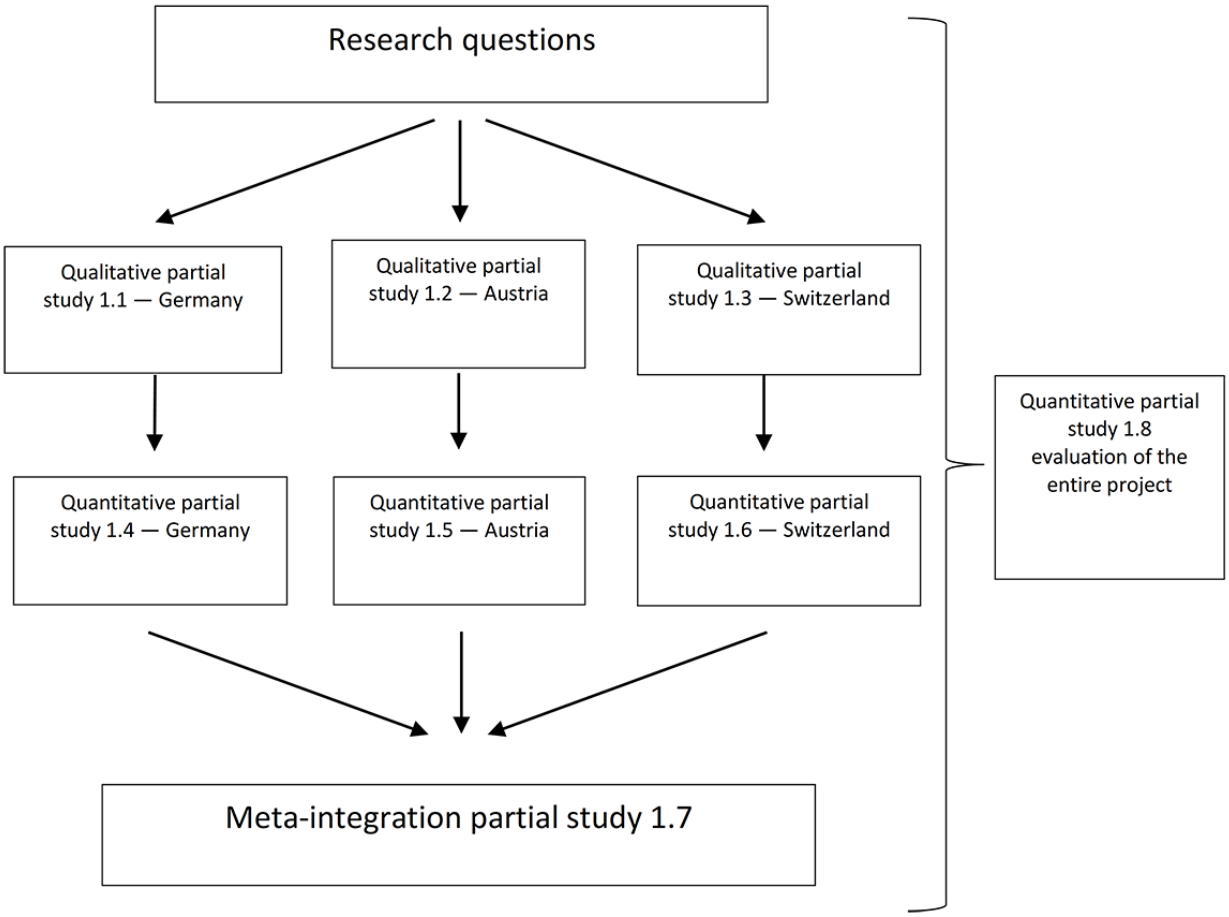
To ensure the reporting quality regarding the research methodology of this qualitative study, COREQ (Consolidated Criteria for Reporting Qualitative Research) was used [11]. A checklist including the COREQ items taken into consideration has been attached to the paper (Multimedia Appendix 1). The registration identifier of the study is the International Registered Report Identifier: PRR2-10.2196/15399.

Ethical Considerations

This study (corresponding to partial study 1.2 in Figure 1), led by DM, has received ethics approval (S-287/2020) from the ethics committee of the University of Witten/Herdecke, Germany. Participants were not compensated for their participation.

XSL•FO
RenderX

Figure 1. Characteristics, Opportunities, and Challenges of Osteopathy project flowchart.



Setting and Sampling

Related to the research topic, osteopaths in Austria were questioned through guideline-based interviews. All the interviewed osteopaths (10/10, 100%) had completed at least 4 years of training as osteopaths and practiced as osteopaths in Austria. Instead of asking individual osteopaths to participate, OEGO was contacted with the study project itself, thus avoiding cold-calling. In this way, the criterion of comprehensive training in osteopathy was fulfilled, because otherwise, the participants could not be members of the professional association. This procedure ensured that the participants had provided evidence of their competence. To obtain the widest range of views, the sample was intended to show a high degree of diversity among the participants. Therefore, a further criterion for the whole group of participants was sex distribution according to the population of osteopathic practitioners in Austria. According to the OEGO's membership register, two-thirds of practicing osteopaths in Austria are female and one-third are male. Moreover, care was taken to ensure that the residences and workplaces of the participants were subject to as wide a geographical distribution as possible, so that district-specific phenomena could be excluded. As osteopathy is not an independent profession in Austria, the participants should include the different occupational groups that practice osteopathy. Inclusion and exclusion criteria were subsequently formulated (Textbox 1).

OEGO forwarded the contact details of 11 osteopaths. Appointments for an interview were made with 8 (73%) of the 11 osteopaths. No appointment could be made with 1 osteopath during the study period. From 1 other osteopath, no response to the request was received. Another 1 osteopath did not want to participate in the study; 2 new osteopaths were suggested by the osteopaths themselves. A total of 10 interviews were thus conducted.

All the participants were contacted via email and received a letter containing information about the course of the study, declaration of consent to participate in the study, and data protection declaration. The entire participation in the research project and the answering of individual questions was on a voluntary basis; nonparticipation did not lead to any disadvantages for the participants. The participants always had the option to end the survey (eg, in the case of unexpected, stressful questions). Through the format of web-based survey, any increased risk of infection for the participants owing to the COVID-19 pandemic could be excluded. Explicit cancellation criteria for the project were not set.

XSL•FO
RenderX

Textbox 1. Inclusion and exclusion criteria.

Inclusion criteria

- The individual has completed 4 years of osteopathic training.
- The individual is currently practicing osteopathy.
- The individual has provided consent to participate in the study.

Exclusion criteria

- The individual cannot be interviewed within the examination period.
- The individual does not have the technical equipment required for participation in a web-based survey.
- The individual demands compensation.

Development of the Interview Guidelines and Data Collection

The guideline-based interview was chosen as a suitable research tool, because the aim of the data collection was to obtain concrete statements about the practice of osteopathy in Austria. In addition, the use of a guideline increased the comparability of the individual data sets [12]. Furthermore, this procedure avoided the possibility that essential aspects of the research question might be overlooked in the interviews [13]. The interview guideline was developed by JM on the basis of the research questions and 2 previous qualitative, partial studies of the COCO project [14,15]. The interview guideline was developed by JM on the basis of the research problem and 2 previous qualitative, partial studies of the COCO project [14,15].

The questions were formulated as open questions (Textbox 2) and arranged according to the groups of topics (Textbox 3). In addition, sociodemographic data about the participants were collected. Alternative questions were prepared to be able to react flexibly to the course of the interview and to respond to the potential needs of the participants. To maintain the flow of the conversation, additional questions were developed in advance. Before beginning the interviews, a test interview was conducted with a German osteopath to test the interview guideline in practice and to improve the interview technique. Finally, the questionnaire was discussed and adapted together with JP, an osteopath with experience in qualitative research.

Textbox 2. Example interview questions.

- Where did you first hear about the osteopathic profession?
- How would you define osteopathy?
- In your opinion, what differentiates osteopathy from other professions?
- What does a typical osteopathic treatment look like for you?
- Should osteopath be its own profession?
- Are there any difficulties or problems that you face in your daily work life as an osteopath?
- Has the corona pandemic changed anything in your daily practice?

XSL•FO
RenderX

Textbox 3. Contents of the interview guideline.

Training and work in osteopathy

- Educational background
- Motivation
- Training structure
- Work experience
- Acquisition structure
- Fields of activity

Characteristics of osteopathy

- Definition
- Properties
- Differentiation of professional profile
- Competences
- Features
- Limits

Challenges of osteopathy

- Health value
- Employment policy
- Obligations
- Restrictions
- Conflicts

Chances and opportunities in osteopathy

- Perspective
- Research
- Desires

Data Collection

Data were collected from November 29, 2021, to January 26, 2022. In total, 10 interviews were conducted. The interviews were conducted using the Zoom software (Zoom Video Communications; audio and video were recorded). Apart from research economy and temporal and local flexibility, interviews were primarily conducted on the web to protect the participants from infection during the COVID-19 pandemic. No other individuals were present during the interviews. No interview was repeated. The interviews were recorded as a video file for transcription. The participants agreed in writing to the archiving of the files until the end of the publication activities or up to a maximum of 5 years after data collection. The files were protected against unauthorized access and stored and evaluated on local data carriers of a password-protected computer. Only encrypted files were transferred among the study colleagues. Upon completion of the study project, the recordings of the interviews shall be deleted irrevocably. All the interviews were conducted by JM in German. The interviewer is male, holds a master of science degree, and has already published in 2019 within the COCO project. At the time of the survey, he worked independently as a physiotherapist in his own practice in Germany and was a doctoral student at the University of Witten/Herdecke. For this research project, JM was trained within a 3-part seminar at the Freie Universität Berlin regarding the collection and evaluation of qualitative data and the conduction of interviews. There was no previous personal relationship with any of the interviewees. The recordings were transcribed by JM. Transcription was performed according to pre-established rules, which were consistently observed, as there are no generally accepted transcription rules [16]. The rules were based on the transcription rules of Kuckartz and Rädiker [17] for computer-assisted evaluation. The participants were sent the transcript to gather their comments, if any, and to receive their final approval.

Data Analysis

On the basis of the results of the interview studies already conducted within the COCO project, deductive (ie, indirectly theory-driven) categories were formed first. It is indirect because the categories are descriptive and their definition is not the basis of a theory-driven description. As a first step, classical deductive codes were derived from the interview guideline (deductive category application). The transcripts were analyzed in the original language by means of content structuring. With the

XSL•FO
RenderX

help of the MAXQDA 2022 software (VERBI Software) [18], quotations that were relevant for the abovementioned research questions were categorized. Owing to an extensive interest in further knowledge, the preselected segments were expanded or specified by means of inductive categorization to deduce further important aspects (deductive-inductive categorization) [17]. The quotations were summarized in categories, which were subsequently grouped into high-level categories. The main strategies of data analysis are described in Textbox 4.

Textbox 4. Data analysis strategies.

Coding

- After developing an initial category system together (first, deductive; second, inductive), 2 complete material iterations were performed by 2 separate evaluators

Discussion

- Comparison and discussion of the results, with the main focus on the integration of different perspectives and the elimination of ambiguities

Quality control

- To check the quality of the category system (intercoder reliability) with its coding rules by means of the Cohen κ coefficient

Final iteration

- After the quality control step, a final, complete material iteration was performed by an evaluator using the final category system

Shared Coding

A first interview was coded by 2 evaluators together (JM and UW), and the category system was inductively expanded. An initial category system was developed, and coding rules were determined on the basis of 2 partial studies (eg, Figure 1—partial studies 1.1 and 1.3) [14,15] and the interview questions and guidelines developed specifically for this partial study. The result was an initial, deductive category system, and the first coding rules were defined. The deductively determined categories were based on the research questions and increased based on inductive subcategories during the evaluation. Then, the first iteration was run with the entire material, followed by further inductive categorization (UW). A second evaluator (JM) ran the second iteration of the entire material based on the category system resulting from the first material iteration. The results were then compared and discussed, with the main focus being on the integration of different perspectives and the elimination of ambiguities. The code definitions and anchor examples were also revised in this step. Finally, the category system was standardized. This shared coding from the beginning of the study was intended to increase the intersubjectivity of the statement.

Quality Control

Subsequently, of the 10 interviews, 3 (30%) were selected via lot procedure as a subsample to check the quality of the category system with its coding rules by means of the Cohen κ coefficient. The determination of the Cohen κ coefficient is a method for checking the intercoder reliability [19]. UW received the 3 allotted coded interviews from JM. The segment boundaries of the encodings were retained during this iteration. In this way, the evaluator was able to recognize which text parts were encoded and arrange these according to the categories. Without random adjustment, a match of 62.1% between the 2 encodings of this subsample was reached. The randomly adjusted coefficient was 0.61. Pursuant to Altman [20], accordance is considered as “good” in the case of a κ coefficient of 61-80. After this quality control, as there was good accordance; a final complete material iteration was performed by an evaluator (JM) using the final category system.

Results

Overview

The result of this study was a system of categories in which the statements of the participants were classified and subcategorized according to the research questions regarding the groups of topics, such as characteristics, challenges, and opportunities of the osteopaths practicing in Austria. Sociodemographic data and more general statements about the practice of osteopathy were also classified in the main category, “training and work in osteopathy.” Out of current concern, the osteopaths were also asked about the influence of the COVID-19 pandemic on their work. On the basis of the interesting statements, we decided to dedicate a special category to this topic.

The final category system consists of 71 categories with a total of 783 encoded text passages. A definition was created for each individual category, and a quotation was recorded as an anchor example.

Only a part of these results could therefore be described in this paper. The printed quotations have been translated into English and serve for illustration purposes only. The question whether individual statements of the participants (O1 to O10) represent the entire population of osteopaths practicing in Austria shall be subject to further investigation.

Training and Work in Osteopathy

A total of 10 osteopaths practicing in Austria were interviewed—6 (60%) women and 4 (40%) men. Of the 10 osteopaths, 3 (30%) were physicians and 7 (70%) were physiotherapists. Of the 10 interviewees, 9 (90%) worked independently in their own practices and 1 (10%) had an employment relationship at the time of the interview. Only 10% (1/10) had other employees. The longest interview lasted 61

XSL•FO
RenderX

minutes, and the shortest interview lasted 29 minutes. The average duration of the interviews was 46 (SD 9.2) minutes. The participants' characteristics are described in Table 1.

Table 1. Participants' characteristics (N=10).

Characteristics	Participants, n (%)
Sex	
Male	4 (40)
Female	6 (60)
Profession	
Physician	3 (30)
Physiotherapist	7 (70)
Degree	
Diploma in osteopathy	3 (30)
Master of science	6 (60)
Nondegree	1 (10)
Training facility	
Vienna School of Osteopathy, Vienna, Austria	8 (80)
European College of Osteopathy, Munich, Germany	1 (10)
The International Academy of Osteopathy, Darmstadt, Germany	1 (10)
Clinical experience (y)	
1-5	2 (20)
6-15	5 (50)
>15	3 (30)

Characteristics of Osteopathy

On the basis of the descriptions provided by the participants regarding the properties of osteopathy, the following categories were developed: "definition of osteopathy," "patient profile," "anchor personalities and literature," and "limits of the treatment technique."

Definition of Osteopathy

When asked about the definition of osteopathy, several participants had difficulties in explaining the concept:

> *Yes, it's really a difficult question.* [O1; item 35]

The explanation of the term was mostly based on the manual work, origin of the word, differentiation or overlapping with other professional groups, or citation of definitions of third parties. Often, reference was made to the philosophy of osteopathy, holism of the treatment method, and activation of self-regulating forces. There was no uniform definition of osteopathy among the answers of the interviewees.

Patient Profile

Most of the interviewed osteopaths treat patients of all ages:

> *Oh, everything actually, there are patients of all ages. From...three-month-old babies to over 90-year-old men, women, so I couldn't paint a typical picture.* [O2; item 54]

Common indications for child treatments mentioned by the interviewed osteopaths are sleep disorders, torticollis, scoliosis, asthma, abdominal colic, and plagiocephaly. The treatment of adults was mainly based on the diagnosis or leading symptoms from the orthopedic area: back and neck pain and joint pain. Neurological diseases such as Parkinson disease were also mentioned. Moreover, patients were regularly treated for headache, migraine, tinnitus, chronic pain, craniomandibular dysfunction, abdominal pain, and hormonal imbalances or if wishing to become pregnant. However, internal diseases such as sinusitis, bronchitis, and cystitis were also treated by the physician O2 on the basis of osteopathic methods.

The indication of the treatment was usually given by the patients' treating physician:

> *Patients are often assigned by the doctors, meaning that the doctor writes a prescription with a recommendation to contact a certain therapist.* [O8; item 51]

A participant working as a general practitioner in addition to his osteopathic activity also acquired patients during his regular consultation as a physician:

> *And actually, many of those who go to the general practitioner's clinic in the village, they come to me, too.* [O2; item 54]

It appears that, in general, a broad medical field was covered by osteopathy. The selection of the appropriate therapist seemed to depend on personal recommendations of others, on the therapeutical possibilities in the patients' vicinity, and on the training or specializations of the osteopaths:

XSL•FO
RenderX

> *Because the patients who come to me come by word of mouth, yes.* [O5; item 47]
>
> *Everything else is, I think, very average, that is, all the people who come to me do so because I am in their vicinity.* [O3; item 38]
>
> *Patients who travel a long way mainly come because of endocrinological, metabolic problems..., gynaecological problems..., that is to say, where...the focus of my...training has mainly been during recent years.* [O3; item 38]

Anchor Personalities and Literature

It was noticeable that many participants referred to other individuals when answering questions about osteopathy in theory and practice. These "anchor personalities" seem to have a great influence on the self-image of osteopaths in Austria and have therefore been included in a separate category. Both historical personalities from the history of osteopathy and currently active osteopaths were repeatedly mentioned:

> *But there will always be people who really care about this innermost quality of osteopathy.... And just as osteopathy has developed from Still to Sutherland, Becker, Viola Frymann...and all their names..., or Mitchell and Jim Jealous now..., so it will continue to develop.* [O5; item 84]

With regard to the self-study of osteopaths, primarily only the German-language journals of osteopathy were mentioned as reading material. Of the 10 participants, only 1 (10%) indicated that they regularly read an English journal:

> *Yes, I regularly read the two journals, DO and Osteopathische Medizin.* [O5; item 88]

Limits of Osteopathy

There was agreement regarding the limits of osteopathy. The primary treatment of structural injuries or the care of patients with cancer without medical supervision were clearly mentioned by the interviewees as limits of osteopathic activity. However, patients were given osteopathic treatment nevertheless:

> *Yes, of course I see the limits in the pathologies that are there. If there is actually...an osteoarthritis that is simply there and will not vanish, osteopathy shall certainly have its limits; one can perhaps relieve the pain, but the osteoarthritis cannot be cured by osteopathy, now can it. I also see limits in some diseases.* [O7; item 56]

O9 differentiated upon request that it is not the diagnosis that is decisive for the objective of treatment when it comes to whether osteopathic treatment is indicated or contraindicated:

> *Well, that depends on the objective.... It all depends totally on the objective. If I say that I want to treat coxarthrosis curatively, I think that we shall soon reach our limits with osteopathy; if we do a control X-ray after six months, we shall see that it is still coxarthrosis.... But if I say that I want to improve the quality of life, I would treat them nevertheless. The question is what the objective is.* [O9; item 41]

O4 brought another aspect to mind:

> *We must not exceed the limits...of our own competence. That is very important. Unfortunately, many colleagues do this by suddenly giving dietary recommendations, by suddenly recommending medicines...or by talking patients out of taking medicines. Especially now when it comes to vaccination.* [O4; item 75]

Challenges of Osteopathy

The challenges mentioned for the field of osteopathy can be classified mainly into the categories, "identity problem," "disagreements within the osteopath community," "research," "training quality management," and "conflicts and difficulties."

Identity Problem

One of the great challenges of osteopathy is its unclear definition and lack of differentiation from other professional groups. As O6 clearly pointed out, osteopathy has an identity problem:

> *I would say that the identity problem is the most important issue...The definition is the most difficult question and no one can answer that. And if we can't answer that and don't deal with it, how can we argue what we are if we ourselves don't know exactly what we are.* [O6; item 49]

Disagreements Within the Osteopathic Community

The identity problem or the problem of missing a uniform professional self-image might be based on disagreements within the osteopathic community described by several participants. Overall, 2 groups can be identified among practicing osteopaths: on the one hand, the structurally working osteopaths and, on the other hand, the cranially or biodynamically working osteopaths:

> *Yes, there is really a gap between biodynamic osteopathy and structural osteopathy.* [O7; item 104]

This conflict might be decisive not only in terms of a common definition but also with respect to a possible recognition of osteopathy in terms of professional policy. O9, who was involved in professional policy, feared that these disputes might even prevent recognition:

> *The problem concerning regulation is - and that's simply the case now and that's also the elephant in the room about which no one is talking -...the problem with regulation has always been cranio.... You cannot say it openly, but it was always the problem of craniosacral therapy, no matter who I talked to.* [O9; item 93]

Those participants who worked biodynamically were more critical toward regulation:

> *If one tries now to take this out of this mental...source,... I see the risk that it is practically shifted into evidence-based, as important as that is, well, but only into evidence-based, visible and perceptible dimensions, then osteopathy shall lose its*

XSL•FO
RenderX

soul from my point of view.... And that's actually the greatest threat to osteopathy for me. [O5; item 64]

Research

In this context, O1 pointed out that evidence-based research can only substantiate a certain part of osteopathy scientifically, whereas other aspects might be lost:

A good scientific basis in order to argue how...many benefits osteopathy has...- in the end,... academisation probably cannot be avoided and will certainly be necessary. Even if all these developments are not entirely without risks. That is, the risk of losing sight of the holistic aspects of osteopathy. [O1; item 81]

Training Quality Management

Many osteopaths considered existing weaknesses in the training courses and their structures as a further obstacle. According to O4, a central aspect is the inadequate teaching of the skills for scientific work at osteopathy institutes and, thus, the lack of evidence-based research in osteopathy:

[Oh my] the training...We should learn from the beginning, not only during the last year when we have to write a Master Thesis, we should learn from the beginning what it means to work in an evidence-based way, to do research...It works in physiotherapy and is continuously getting better there, but in osteopathy...At the beginning, we never learn to deal with available studies, it is a matter of training. From the beginning, not only during the fifth year shortly before the Master Thesis, we should have the first lessons in statistics. [O4; item 113]

Conflicts and Difficulties

The lack of clarity regarding the profession is evident in the differentiation with respect to other professional groups. With regard to the settlement for osteopathic treatment with health insurance providers, several participants also reported potential for conflicts. Osteopathic treatment is often provided on the grounds of a physiotherapeutic prescription by the physician. The reasons are the economic pressure on the practices or the social situation of the patients:

Many colleagues work as physiotherapists, they also charge for osteopathy as physiotherapy, and yes, they are refunded in this way. And as a result,...osteopathy is also a little...less in the focus than it should be. I've been working for 20 years now, I'm only writing osteopathic invoices.... But...of course...I understand the problem. If somebody has fewer patients...and...has to charge for...physiotherapy, I absolutely understand the situation.... But...these problems are of course...long-burning issues. [O5; item 78]

No, [the bill] of course says physiotherapy and remedial massage, because otherwise the patient doesn't get his/her money from the insurance company. For the insurance company, well, this is ok or it is tolerated. I have already received the feedback from many patients that they told the company that they went to an osteopath, and the health insurance said that of course that can't be billed, [but] we shall write physiotherapy and remedial massage and then that's it. [O8; item 55]

When asked about the challenges for osteopaths in general, the physicians working as osteopaths did not report any difficulties related to their practice. A physician and osteopath, in contrast, was aware of the potential for conflict:

Yes, of course I know that. I have a bonus, because I'm simply a doctor. And of course, osteopaths that aren't doctors have greater difficulties and are often rejected,...well,...because they are no medical doctors in a manner of speaking and...there are obviously difficulties. [O1; item 101]

Another participant stated it even more clearly:

No, I'm a doctor, I have...no restrictions. [O9; item 85]

Opportunities in Osteopathy

Regarding the questions about the opportunities and chances in osteopathy, most of the statements could be classified into the categories "professional profile" and "position in the health care system." A central opportunity in osteopathy is the installation of an independent profession. Almost all the osteopaths explicitly formulated the desire for their own professional profile. However, there was disagreement about the questions regarding where and how this profession should be integrated into the health care system or which competences it should include:

In the midst of the other health professions..., well, I don't see us as special consultants, as it is now in America, for example. But I see us as a health profession next to physiotherapists, occupational therapists... [O4; item 81]

In this context, many possible applications were mentioned for the field of osteopathy. An osteopath saw a great opportunity in the prevention of diseases:

Concerning also prevention..., I believe that osteopathy has an enormous potential for people's health by simply doing something really good and also really preventing things,...follow-up problems or operations or God knows what...I see a huge opportunity there. [O10; item 107]

O2 also attributed the potential for cost reduction to preventive osteopathy. From his point of view, examinations and medical consultations might be reduced:

I see a huge and very central importance of osteopathy in primary care...and I am convinced...from my daily experience that an incredible number...of diagnostic measures or specialist care...might be avoided if people were primarily also treated by osteopaths. [O2; item 68]

Whether osteopathy actually contributes to disease prevention and can thus also lead to cost reduction or relief for the health care system is to be investigated using clinical study designs

XSL•FO
RenderX

on the effectiveness of the treatment method itself. The position of osteopathy in the health care system and its differentiation from other professional groups have also not been uniformly described by practitioners in other countries.

The COVID-19 Pandemic

For current reasons, the participants were questioned about the effects of the COVID-19 pandemic, existing since March 2020, on their professional activities. Similar to a magnifying glass, crises very often reveal the weaknesses and failures of structures and concepts; however, they can also show their viability and strengths. Most respondents described the time of the COVID-19 pandemic with the lockdowns and the associated measures as a turning point in their practice. However, none of the participants described economic losses or existential fear:

> *Well, the time during Covid-19 wasn't easy at all.* [O5; item 75]

At the time of the survey, everybody had to wear a face mask, patients and osteopaths alike. The interviewees not only described the difficult communication with the patients because of the mask but also mentioned limitations during examination and treatment. Certain treatments, for example, techniques relating to the mandibular joint, could not be performed for patients wearing a face mask:

> *I believe that a lot of communication is lost through the mask, because you don't see the whole face of the patient. Of course, you have communication through the eyes, but there is still a barrier, a lot is lost...It already starts with and continues during inspection: you only see half of the face and in the case of jaw problems, I have to take down the mask first.* [O8; item 107]

Almost all the interviewees described a change in the clientele of their patients. Stress, sleep disorders, headaches, and dysfunctions of the mandibular joint were increasingly mentioned:

> *Psychosocial stress is increasing immensely...This in turn results...in sleep disorders...Mandibular joint problems due to stress, but also - and this is my own observation - because you constantly want to push around this mask if you have to wear it all day...I think that this has a huge influence...[O6; item 79]*

Teaching also seemed to be affected by the protection measures for pandemic control. An osteopath reported that teaching on inpatients at the hospital ceased. The question of whether the osteopathic treatment of inpatients in institutions was disturbed by a lack of external osteopaths remained unanswered:

> *Prior to the lockdown, our osteopathic child centre also paid visits to the neonatal ward...where we treated premature babies. Unfortunately, this is not possible at the moment.* [O1; item 47]

Owing to the cancellation of congresses and courses or their transfer to the digital world, interviewees experienced a gap in their personal training plans:

> *Of course, I...repeatedly attended courses. However, I scarcely did so in the last two years, actually...[O5; item 88]*

This can only be a small insight into the impact of the pandemic in the field of osteopathy. The effects of the COVID-19 pandemic on osteopathic care should be investigated systematically in the next few years.

Discussion

Principal Findings

This study identified numerous aspects, possibilities, and opportunities in osteopathy in Austria from the point of view of the osteopaths practicing in Austria.

In our survey, the typical osteopath presents as female and has previously worked as a physiotherapist, as previous studies have found [8]. This is consistent with other surveys from Europe regarding osteopathy. Moreover, in accordance with a study from Italy, the typical osteopath practicing in Austria works independently in their own practice and without employees [21].

The osteopaths interviewed usually found it difficult to define osteopathy. The respondents were not able to provide a uniform definition of osteopathy. Many respondents even expressed difficulties in precisely describing their profession. Nevertheless, recurring patterns can be recognized in the explanations given by the respondents.

The participants attempted to define osteopathy by drawing a distinction or differentiation from other professions and using third-party definitions. Furthermore, the philosophy of osteopathy, various osteopathic concepts or models of thought, the holistic nature of the treatment method, and activation of the patient's self-healing powers are often referred to. A possible reason for the heterogeneous attempts at explanation may lie in the difference in training and previous education. A recent study showed that only 17% of osteopaths surveyed in Austria identified themselves "exclusively" as osteopaths [8]. Therefore, there is a suspicion that, as our study also showed, the basic profession and nonregulation have a major influence on self-image. We observed a fundamental distinction between therapeutic and medical osteopaths.

From our point of view, the clear statements regarding the disagreements within the osteopath community were surprising. The conflicts do not remain in the specialist circles of osteopathy, but they even extend to the level of professional policy. The question is whether this is a country-specific observation for Austria. In their study in Australia in 2018, for example, Blaich et al [22] found disagreement about the specialization of osteopaths; however, it did not result in the splitting of osteopaths into 2 separate groups. The belief patterns and paradigms of individual treatment techniques that influence professional identity are not new in osteopathy [23]. The fact that, according to the osteopaths interviewed, these intraprofessional conflicts exist even on the political level or are the reason for nonregulation is remarkable. An increasing number of European countries regulate the professional practice of osteopathy. Therefore, it remains to be investigated whether this dispute itself has an influence on the nonregulation of

XSL•FO
RenderX

osteopathy in Austria. However, conflicts and different opinions within a professional group are not inherent in osteopathy; these also exist in other medical professions such as chiropractic [24,25].

The general development of a profession is not only subject to cultural, historical, and social influences but also to the question of gender [26]. In this context, this study indicates a large influence of anchor personalities on the self-image of the interviewed osteopaths. It is remarkable that the anchor personalities mentioned are almost exclusively men. The historical context is worth noticing here, because Andrew Taylor Still, the founder of osteopathy, explicitly promoted equality between men and women already in the 19th century, in contrast to many other universities or teaching institutes during that time. He expressly included women in his courses [27]. In this context, it should be noted that in other health professions, although the practitioners are predominantly women, the leadership positions are often mainly occupied by men [28]. It is therefore not surprising that most users and practitioners of alternative medicine are women if their health needs are not being met by scientific medicine [29]. This becomes problematic when these professions are or become patriarchally dominated to match scientific standards [30].

To answer the questions about the origin of these conflicts and to deal with these in the future, we believe that a systematic and country-specific scientific analysis will be required. Conflicts in the health care system not only have the potential to weaken a profession but can also have a stimulating influence if understood as an opportunity [31].

Most of the osteopaths surveyed were in favor of a legal regulation of the profession. Under certain circumstances, osteopathically trained physiotherapists could benefit more from this, as they currently still need a physician's order to be able to practice with legal certainty.

However, nonregulation also has also some advantages—no applications for licenses, no obligation for regular further training, and unregulated pricing for treatment. With integration into the health care system, some participants fear deterioration owing to possible low or lower payment by health insurance companies.

Training quality management and studies in the subarea of osteopathy were also mentioned as challenges in this context. In Italy, Sweden, and Australia, the transfer of scientific results to the practical work of osteopaths has already been systematically investigated in a country-specific manner [32-34]. The openness to evidence-based practice (EBP) appears to exist among practicing osteopaths on a transnational basis, but the skills in dealing with the former vary from country to country. A study of EBP from Spain characterized the skills of the participants to deal with EBP as being rather low. This might be related to the lack of legal regulations and the inadequate transfer of knowledge in the training institutions [35]. The situation regarding osteopathy is similar in Austria. Additional country-specific studies are required to identify conclusions and connections.

The different situation in everyday practice owing to the COVID-19 pandemic and the respective infection protection measures also had an influence on the daily work of osteopaths. Several interviewees realized an evident change in the patients' profile. Although economic damage or fear for their professional existence were not explicitly described, most osteopaths working independently were themselves responsible for the implementation of the legal measures in their practices. The impact of the pandemic on the daily work in practice seems to have been less considerable than the impact on the field of training in osteopathy. As a large part of practical teaching occurs with patients under supervision, it is difficult to implement in a web-based format. The impact of the pandemic on clinical research at universities or universities of applied sciences remains to be examined.

In the case of further investigations in this area, we recommend a specific distinction of the participants between physicians and physiotherapists practicing as osteopaths. As there is no uniform training or legal regulation of osteopathy in Austria, only physicians and physiotherapists trained in osteopathy exclusively practice osteopathy. The results of this study suggest that there are evident differences between these 2 professional groups regarding, for example, patient acquisition, conflict management, and cooperation with other professional groups.

The extent to which the individual statements made by the interviewees represent the entirety of osteopaths practicing in Austria will be further investigated. The protocol of the COCO project describes the further procedures. The results of the qualitative partial studies (studies 1.1, 1.2, and 1.3 in Figure 1) will be combined in a following study to verify the results of the qualitative partial studies in relation to the population [36]. We will develop a standardized questionnaire as a measuring instrument.

An important feature of this study is the methodology, including 2 evaluators who completed the entire evaluation process. Through this approach, intersubjectivity increased and new, inductively formed categories were created. During this phase, many aspects of the research problems could be identified and categorized. The intercoder reliability was tested and found to be viable within this study. With another material iteration, the category system can be further refined, and the intercoder reliability can be further increased by optimized code definitions.

Limitations

First, it should be noted that the results of this study do not necessarily allow conclusions to be drawn about the entirety of osteopaths in Austria, as this is not an evaluation of representative surveys with large numbers of participants. Nevertheless, certain tendencies seem to emerge when statements by osteopaths appear to be congruent, that is, confirm each other or complement each other in a meaningful way. The sample represents the entirety of osteopaths in Austria well. Most respondents were women and physiotherapists [8]. Nevertheless, bias cannot be dismissed with such a specific sample. However, they give an idea about how osteopaths in Austria think, and the results obtained can serve as a hypothesis for large quantitative studies to test.

XSL•FO
RenderX

Conclusions

It is difficult to characterize the community of osteopaths in Austria conclusively. On the one hand, there is a great deal of agreement about the urgency regarding regulatory legislation for their profession, a necessary revision of training structures, and the specific promotion of scientific studies of osteopathy. However, when it comes to the concrete practice of osteopathy, deep trenches and even strong disputes have occurred among osteopaths.

The following question remains to be answered: what is "correct" or "true" osteopathy? If we consider that osteopathy has derived from various sources; that its founder did not give a final answer to the question about what he understood by osteopathy; and that each discipline is constantly developing, solely through the different osteopaths practicing, it appears that this question cannot be answered completely.

Apart from this issue, there is another and equally sensitive question, that is, whether and how the different parties can or even must be brought together for the regulation of their profession, which is desired by most of them. The different professional origins of osteopaths should also be considered. With regard to binding legal regulations, which would not least strengthen the professional image, mutual understanding seems to be imperative. Perhaps such an understanding might also lead to greater political weight for osteopathy, which it urgently needs, not only in terms of legal regulations but also to be able to promote important research projects.

The question arises as to whether the conflicts within osteopathy, in particular, with their possible professional-political consequences and the immense influence of the basic profession in the practice of osteopathy, are a country-specific phenomenon for Austria. However, there is a lack of studies in German-speaking countries with comparable qualitative designs to assess the work of osteopaths in more detail. We are therefore planning a meta-synthesis of qualitative studies with the aim of generating new theoretical insights from the accumulation of study results. Both the studies from the COCO project itself and other relevant literature can be used for the meta-synthesis.

To the best of our knowledge, the COCO project is the largest mixed methods study project on the osteopathic profession in German-speaking countries. The category system with its reliability check can be used as a basis for a repetition of the study. Such a research project would also be interesting if the profession was regulated formally and substantially in the near future. The results presented in this paper are not only intended to serve as a basis for further studies but also to provide universities, schools, professional associations, and politicians with an insight into the situation of osteopaths in Austria.

Conflicts of Interest

None declared.

Multimedia Appendix 1
COREQ (Consolidated Criteria for Reporting Qualitative Research) checklist.
[PDF File (Adobe PDF File), 870 KB - humanfactors_v11i1e45302_app1.pdf]

References

1. Benchmarks for training in traditional / complementary and alternative medicine: benchmarks for training in osteopathy. World Health Organization. 2010. URL: https://apps.who.int/iris/handle/10665/44356 [accessed 2023-08-19]
2. Osteopathic Education – Osteopathic International Alliance. URL: https://oialliance.org/resources/osteopathic-education/ [accessed 2023-12-24]
3. Regulation of osteopathy in Europe. European Federation & Forum for Osteopathy. URL: https://osteopathyeurope.org/regulation-2/ [accessed 2023-11-19]
4. Rechtliche situation. Österreichische Gesellschaft für Osteopathie (OEGO). URL: https://www.oego.org/home/alle-infos-uber-osteopathie/ [accessed 2023-11-19]
5. Phillips AR. Professional identity in osteopathy: a scoping review of peer-reviewed primary osteopathic research. Int J Osteopath Med 2022 Sep;45:25-37. [doi: 10.1016/j.ijosm.2022.06.005]
6. Krönke K. A questionnaire to evaluate the professional field of osteopathy in Austria pilot study 2003. Osteopathic Research Web. 2006. URL: https://www.osteopathicresearch.org/files/original/b4b7e15719626458f7ddea5b0dfea05004b992d8.pdf [accessed 2023-12-14]
7. Graf B. OsteopathInnen in Österreich Bestandsaufnahme, Vergleich mit internationalen Daten und Untersuchung der Entwicklung des Berufsbildes und -standes seit 2003 [Master-Thesis]. Krems, Vienna: Donau Universität; 2011. URL: http://tinyurl.com/26uepnmn
8. van Dun PL, Arcuri L, Verbeeck J, Esteves JE, Cerritelli F. The Austrian osteopathic practitioners estimates and RAtes (OPERA): a cross-sectional survey. PLoS One 2022 Nov 28;17(11):e0278041 [FREE Full text] [doi: 10.1371/journal.pone.0278041] [Medline: 36441692]
9. Mayring P. Qualitative Inhaltsanalyse: Grundlagen und Techniken. 12th edition. Weinheim, Germany: Beltz & Gelberg; 2015.
10. Kuckartz U. Mixed Methods: Methodologie, Forschungsdesigns und Analyseverfahren. Berlin, Germany: Springer-Verlag; 2014.

XSL•FO
RenderX

11. Tong A, Sainsbury P, Craig J. Consolidated criteria for reporting qualitative research (COREQ): a 32-item checklist for interviews and focus groups. Int J Qual Health Care 2007 Dec;19(6):349-357. [doi: 10.1093/intqhc/mzm042] [Medline: 17872937]
12. Deppermann A. Interviews as text vs. interviews as social interaction. Forum Qual Sozialforschung Forum Qual Soc Res 2013;14(3):2064 [FREE Full text] [doi: 10.17169/fqs-14.3.2064]
13. Flick U. Qualitative Forschung: Theorie, Methoden, Anwendung in Psychologie und Sozialwissenschaften. Berlin, Germany: Rowohlt Taschenbuch; 1999.
14. Meier MD, Porthun J. Situation, Chancen und Herausforderungen der Osteopathie in der Schweiz. DO Dtsch Z Osteopath 2020 Sep 15;18(04):26-31. [doi: 10.1055/a-1205-2394]
15. Bücker AK. Eigenschaften, Chancen und Herausforderungen der Osteopathie aus Sicht der Osteopathen in der Schweiz [Master-Thesis]. Dresden, Germany: Dresden International University; 2020.
16. Gläser J, Laudel G. Experteninterviews und qualitative Inhaltsanalyse: als Instrumente rekonstruierender Untersuchungen. 4th edition. Berlin, Germany: VS Verlag für Sozialwissenschaften; 2010.
17. Kuckartz U, Rädiker S. Qualitative Inhaltsanalyse. Methoden, Praxis, Computerunterstützung. Weinheim, Germany: Beltz Juventa; 2022.
18. VERBI software. MAXQDA. 2022. URL: https://www.maxqda.com/about [accessed 2023-12-12]
19. Cohen J. A coefficient of agreement for nominal scales. Educ Psychol Meas 2016 Jul 02;20(1):37-46 [FREE Full text] [doi: 10.1177/001316446002000104]
20. Altman DG. Practical Statistics for Medical Research. New York, NY: Chapman and Hall/CRC; 1991.
21. Cerritelli F, van Dun PL, Esteves JE, Consorti G, Sciomachen P, Lacorte E, OPERA-IT Group. The Italian Osteopathic Practitioners Estimates and RAtes (OPERA) study: a cross sectional survey. PLoS One 2019 Jan 25;14(1):e0211353 [FREE Full text] [doi: 10.1371/journal.pone.0211353] [Medline: 30682169]
22. Blaich R, Steel A, Clark D, Adams J. Challenges and opportunities for Australian osteopathy: a qualitative study of the perceptions of registered osteopaths. Int J Osteopath Med 2018 Dec;30:18-24. [doi: 10.1016/j.ijosm.2018.10.004]
23. Grace S, Fleischmann M, Vaughan B. "If you dont use or understand visceral osteopathy you are not a real osteopath": professional identity in Australian osteopathy through the lens of a single traditional technique. Explore (NY) 2021 Nov;17(6):535-540. [doi: 10.1016/j.explore.2020.07.001] [Medline: 32736952]
24. Pollard H. Reframing a debate in chiropractic. Chiropr Man Therap 2021 Nov 03;29(1):44 [FREE Full text] [doi: 10.1186/s12998-021-00401-5] [Medline: 34732222]
25. McGregor M, Puhl AA, Reinhart C, Injeyan HS, Soave D. Differentiating intraprofessional attitudes toward paradigms in health care delivery among chiropractic factions: results from a randomly sampled survey. BMC Complement Altern Med 2014 Feb 10;14(1):51 [FREE Full text] [doi: 10.1186/1472-6882-14-51] [Medline: 24512507]
26. Hall P. Interprofessional teamwork: professional cultures as barriers. J Interprof Care 2005 May 06;19 Suppl 1(sup1):188-196. [doi: 10.1080/13561820500081745] [Medline: 16096155]
27. Still AT, Hartmann C, MacGovern J. Das große Still-Kompendium: Autobiografie, Philosophie der Osteopathie, Philosophie und mechanische Prinzipien der Osteopathie, Forschung und Praxis. Berlin, Germany: Jolandos; 2013.
28. Feldman DE, Bellavance C, Frédérick D, Gagnon T, Lalonde C, Hudon A. Gender issues in physiotherapy in Quebec, Canada. Physiother Can 2023 May 01;75(2):169-176. [doi: 10.3138/ptc-2020-0142] [Medline: 37736386]
29. Shahvisi A. Medicine is patriarchal, but alternative medicine is not the answer. J Bioeth Inq 2019 Mar 20;16(1):99-112 [FREE Full text] [doi: 10.1007/s11673-018-9890-5] [Medline: 30570716]
30. Flesch H. Silent voices: women, complementary medicine, and the co-optation of change. Complement Ther Clin Pract 2007 Aug;13(3):166-173. [doi: 10.1016/j.ctcp.2007.03.005] [Medline: 17631259]
31. Piryani RM, Piryani S. Conflict management in healthcare. J Nepal Health Res Counc 2019 Jan 28;16(41):481-482. [doi: 10.33314/jnhrc.v16i41.1703]
32. Leach MJ, Shaw R, Austin P, Fryer G, Thomson OP, Adams J, et al. Attitudes, skills, and use of evidence-based practice: a cross-sectional survey of Swedish osteopaths. Int J Osteopath Med 2020 Dec;38:41-49 [FREE Full text] [doi: 10.1016/j.ijosm.2020.10.006]
33. Cerritelli F, Iacopini A, Galli M, Thomson OP, Sundberg T, Leach MJ, et al. Evidence-based practice among Italian osteopaths: a national cross-sectional survey. BMC Complement Med Ther 2021 Oct 07;21(1):252 [FREE Full text] [doi: 10.1186/s12906-021-03430-y] [Medline: 34620142]
34. Leach MJ, Sundberg T, Fryer G, Austin P, Thomson OP, Adams J. An investigation of Australian osteopaths attitudes, skills and utilisation of evidence-based practice: a national cross-sectional survey. BMC Health Serv Res 2019 Jul 17;19(1):498 [FREE Full text] [doi: 10.1186/s12913-019-4329-1] [Medline: 31315615]
35. Alvarez G, Justribo C, Sundberg T, Thomson OP, Leach MJ. A national cross-sectional survey of the attitudes, skills and use of evidence-based practice amongst Spanish osteopaths. BMC Health Serv Res 2021 Feb 10;21(1):130 [FREE Full text] [doi: 10.1186/s12913-021-06128-6] [Medline: 33563266]
36. Porthun J, Manschel J. Characteristics, Opportunities, and Challenges of Osteopathy (COCO) in the perceptions of osteopaths in Germany, Austria, and Switzerland: protocol for a comprehensive mixed methods study. JMIR Res Protoc 2019 Dec 18;8(12):e15399 [FREE Full text] [doi: 10.2196/15399] [Medline: 31850852]

XSL•FO
RenderX

Abbreviations

COCO: Characteristics, Opportunities, and Challenges of Osteopathy
COREQ: Consolidated Criteria for Reporting Qualitative Research
EBP: evidence-based practice
OEGO: Austrian Society for Osteopathy (Österreichische Gesellschaft für Osteopathie)

Edited by A Kushniruk; submitted 23.12.22; peer-reviewed by D Skinner, W Perveen, C Mastronardo; comments to author 20.11.23; revised version received 27.11.23; accepted 27.11.23; published 17.01.24.

Please cite as:
Manschel J, Porthun J, Winkler U, Beuckels JMAT, Martin D
Characteristics, Opportunities, and Challenges of Osteopathy Based on the Perceptions of Osteopaths in Austria: Qualitative Interview Study
JMIR Hum Factors 2024;11:e45302
URL: https://humanfactors.jmir.org/2024/1/e45302
doi:10.2196/45302
PMID:38231542

©Jonas Manschel, Jan Porthun, Ulrike Winkler, Jean Marie A T Beuckels, David Martin. Originally published in JMIR Human Factors (https://humanfactors.jmir.org), 17.01.2024. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on https://humanfactors.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

A Digital Behavioral Activation Intervention (JuNEX) for Pregnant Women With Subclinical Depression Symptoms: Explorative Co-Design Study

Elisa Mancinelli[1,2], BSc, MSc; Silvia Gabrielli[2], PhD; Silvia Salcuni[1], PhD

[1]Department of Developmental and Socialization Psychology, University of Padova, Padova, Italy
[2]Fondazione Bruno Kessler, Trento, Italy

Corresponding Author:
Elisa Mancinelli, BSc, MSc
Department of Developmental and Socialization Psychology
University of Padova
Via Venezia 8
Padova, 35131
Italy
Phone: 39 3342799698
Email: elisa.mancinelli@phd.unipd.it

Abstract

Background: Digital interventions are gaining increasing interest due to their structured nature, ready availability, and self-administered capabilities. Perinatal women have expressed a desire for such interventions. In this regard, behavioral activation interventions may be particularly suitable for digital administration.

Objective: This study aims to exploratorily investigate and compare the feasibility of the internet-based self-help guided versus unguided version of the Brief Behavioral Activation Treatment for Depression-Revised, an empirically supported in-person behavioral activation protocol, targeting pregnant women with subclinical depression symptoms. A user-centered design is used, whereby data are collected with the intent of evaluating how to adjust the intervention in line with pregnant women's needs. Usability and user engagement were evaluated.

Methods: A total of 11 Italian pregnant women with subclinical depressive symptoms based on the Patient Health Questionnaire-9 (scoring<15) participated in this study; of them, 6 (55%) women were randomly assigned to the guided group (age: mean 32.17, SD 4.36 years) and 5 (45%) to the unguided group (age: mean 31, SD 4.95 years). The Moodle platform was used to deliver the interventions in an e-learning format. It consisted of 6 core modules and 3 optional modules; the latter aimed at revising the content of the former. In the guided group, each woman had weekly chats with their assigned human guide to support them in the homework revisions. The intervention content included text, pictures, and videos. Semistructured interviews were conducted, and descriptive statistics were analyzed.

Results: Collectively, the data suggest that the guided intervention was better accepted than the unguided one. However, the high rates of dropout (at T6: guided group: 3/6, 50%; unguided: 4/5, 80%) suggest that a digital replica of Behavioral Activation Treatment for Depression-Revised may not be feasible in an e-learning format. The reduced usability of the platform used was reported, and homework was perceived as too time-consuming and effort-intensive. Moreover, the 6 core modules were deemed sufficient for the intervention's goals, suggesting that the 3 optional modules could be eliminated. Nevertheless, participants from both groups expressed satisfaction with the content and found it relevant to their pregnancy experiences.

Conclusions: Overall, the findings have emphasized both the intervention's merits and shortcomings. Results highlight the unsuitability of replicating an in-person protocol digitally as well as of the use of nonprofessional tools for the implementation of self-help interventions, ultimately making the intervention not feasible. Pregnant women have nonetheless expressed a desire to receive psychological support and commented on the possibilities of digital psychosocial supports, particularly those that are app-based. The information collected and the issues identified here are important to guide the development and co-design of a more refined platform for the intervention deployment and to tailor the intervention's content to pregnant women's needs.

(JMIR Hum Factors 2024;11:e50098) doi:10.2196/50098

XSL•FO
RenderX

KEYWORDS

digital intervention; behavioral activation; feasibility; pregnancy; subclinical depression symptoms

Introduction

Background

Peripartum depression refers to an episode of depression that meets the criteria for persistent or major depressive disorder, with onset occurring during the peripartum period [1]; this definition highlights the direct link between the development of the depressive condition and the bodily changes and overall characteristics inherent to the perinatal period [2]. A distinction should also be made with regard to antenatal and postnatal depression, since antenatal depression is recognized as one of the main predictors of postnatal depression, with the latter then aggravating the repercussions on the mother as well as on the child and the whole family [3]. In Europe, antenatal depression counts a mean prevalence of 17.9% [4], while postnatal depression ranges from an average of 12.91% to 16.62% [5]. In Italy, specifically, the literature highlights a prevalence ranging from 6% to 22% for antenatal depression [6-9] and from around 13% to 23% for postnatal depression [5,7,9,10]. Such percentages emphasize the necessity for early detection and the implementation of prevention programs to alleviate depression symptoms already during pregnancy. Notably, women have expressed a desire for perinatal support programs and have reported benefiting from them, both in terms of symptom reduction and increased sense of agency [11]. However, barriers to help seeking, in the context of perinatal care, have been widely recognized (*knowledge barriers*, eg, difficulty in recognizing health needs and distinguishing emotional difficulties as well as not knowing the services available; *practical barriers*, eg, time and economic constraints; and *attitudinal barriers*, eg, stigma, guilt) [11-13], contributing to the still limited availability of the services to support perinatal women's mental health [14].

Within this context, digital solutions might be particularly valuable. A recent review that specifically focused on the application of eHealth in perinatal care [15] highlighted the potential of these solutions as alternatives or supplements to standard mental health practices, both for screening and intervention. A subsequent review [16] also emphasized the beneficial role of digital solutions in addressing perinatal depression by enhancing accessibility to psychological interventions, thus promoting scalability, which could ultimately allow to work around the abovementioned barriers to help seeking. Digital interventions could thus be valuable solutions to fill in the gap between what is asked and desired by women and the logistic and economic limits on both clinical professionals and health institutions part. Notwithstanding, despite the increasing focus on peripartum depression, there is currently a scarcity of digital psychological interventions aimed at alleviating depression symptoms, particularly those grounded in empirically validated intervention protocols [16-18]. Furthermore, the prevailing focus appears to lean more toward treating rather than preventing perinatal depression. This is evident in the dearth of studies investigating digital interventions during pregnancy and exemplified by the lack of studies investigating digital interventions deployed during pregnancy and to women with subclinical depression symptoms [17,18]. In light of this, there is a need to develop theoretically grounded digital interventions tailored to pregnant women with subclinical symptoms needs and characteristics, ultimately preventing the development or worsening of clinically relevant depression symptoms during the postpartum.

Evidence-based interventions that are brief and structured such as behavioral activation (BA) interventions might be especially helpful to this end, as providing pregnant women with concrete strategies will be useful to support their adjustment. BA is an empirically supported behavioral intervention created to lessen depression symptoms [19-22]; it is based on the idea that a greater awareness of the mutual influence between behavior and emotion can ultimately encourage behavioral change by increasing participation in joyful and adaptive activities while reducing participation in maladaptive behaviors that maintain or exacerbate the depressive symptoms [19,23]. However, a recent scoping review [18] highlighted a gap in the literature: there are few digital BA interventions available during the perinatal period, and none have been specifically deployed during pregnancy. Furthermore, their usability has only been marginally evaluated.

Usability refers to the quality of the interaction occurring between the user (eg, pregnant women) and the tool used (eg, website, smartphone app, etc) [24]; a subcomponent of usability that is more specific is the *user engagement*, which includes the user's cognitive, behavioral, and affective reaction to the tool [25]. These factors are instrumental in supporting user compliance and adherence, and they should be carefully considered and addressed in the development of feasible and acceptable digital interventions [26]. The limited evaluation of these factors in the context of digital mental health solutions may be attributed to the novelty of the field, which has yet to establish a comprehensive understanding of design methods for such tools [27]. When designing these digital solutions, four components should be kept in mind: (1) the design issue and solution, (2) the context in which the design occurs, (3) the dynamics and organization of the design activity, and (4) the actors contributing to the design [28-30]. However, a recent review [27] investigating the design methods and approaches used for the design and development of digital tools for mental health stressed that *human-centered design methods* (ie, the design of digital tools not considering the engineering design and including user-centered approaches, co-design, participatory design, etc) are not yet fully integrated within the field and that reported design approaches are still mainly external, thus excluding the perspective of those for whom the tool is created.

This Study

This study aimed to investigate the feasibility of the Brief Behavioral Activation Treatment for Depression-Revised (BATD-R) [31] protocol that was structured as an internet-based self-help intervention and deployed to pregnant women with subclinical depression symptoms. Compared with other BA

XSL•FO
RenderX

protocols, the BATD-R protocol specifically targets subclinical depression symptoms, is flexible (in terms of both its structuring and the population it is administered to), and can be self-administered or deployed by both specialists or nonspecialists [31].

Given that no previous study has used the BATD-R protocol for this purpose, the intervention developed and evaluated in this study serves as digital "replica" of the in-person BATD-R protocol. By adopting a *user-centered* approach, this study not only aimed to assess its initial feasibility but also sought to gather valuable feedback directly from pregnant women. This feedback will guide the adjustment of the intervention's content and thus its structure, without making assumptions beforehand about the changes required. Indeed, a user-centered design "is an approach to product development that grounds the process in information about the people who will ultimately use the product" [32]; as such, to create a well-accepted, engaging, and effective digital intervention in perinatal care, subsuming the intervention content and the mean through which it is deployed, pregnant women should be consulted in each stage of the intervention's creation and refinement, thereby ultimately allowing the co-design of the final intervention. In this regard, this study relied on the Obesity-Related Behavioral Intervention Trials model [33], which provides an iterative progressive framework guiding the development, testing, and refinement of the behavioral intervention. More specifically, it uses a user-centered design that relies on a data-driven approach to iteratively test and revise the intervention, up until it is deemed appropriate to move to further phases of development and testing, thereby going from the intervention design, its preliminary testing to investigating its efficacy and effectiveness [33].

As previously reported, no digital BA intervention targeting subclinical or clinical depression symptoms among pregnant women has been developed [18]. Nonetheless, it is worth noting that among the existing digital BA interventions, many are guided interventions [18]. The guides, most often mental health specialists or trained professionals, provide additional support to women throughout the intervention, by addressing concerns or supporting them on intervention-related tasks. Mindful of this, a further aim of this study was also to explore and evaluate the role and potential benefits of including a guide as additional support in the self-help intervention. As such, the study also aimed to compare the feasibility of the guided versus unguided version of the intervention.

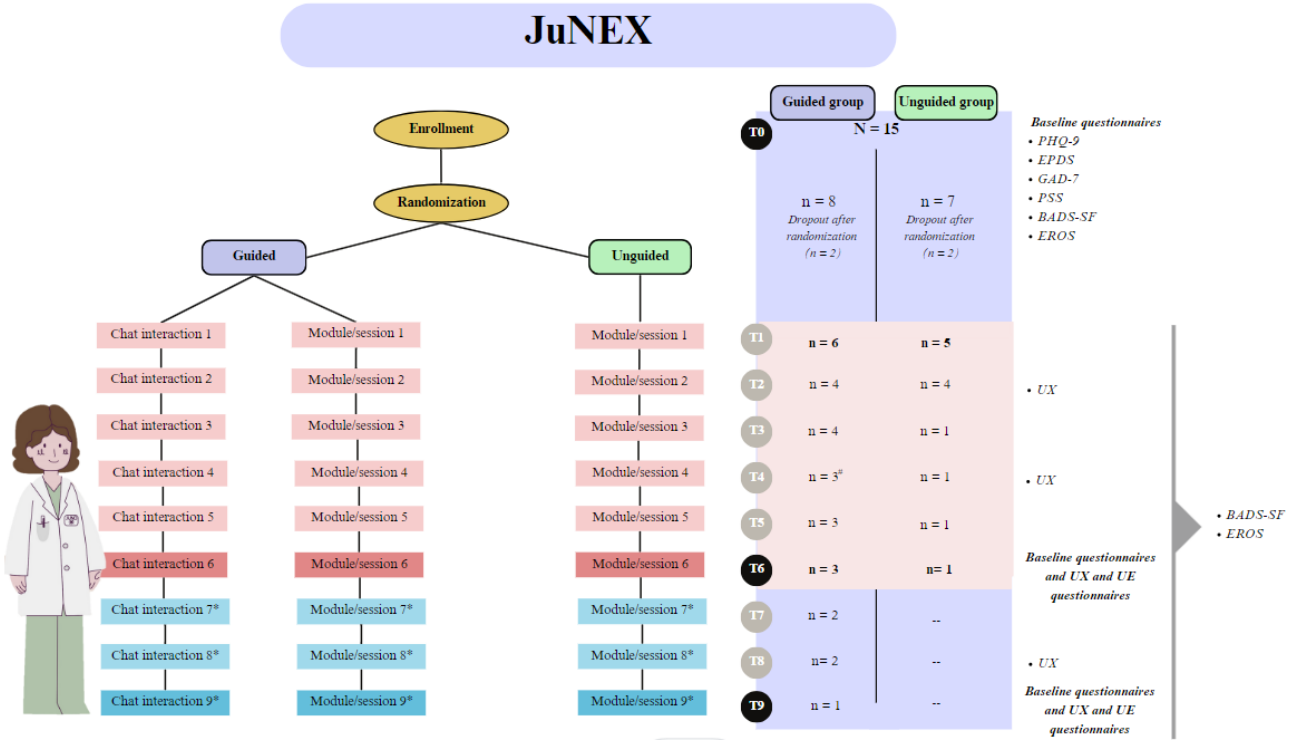
Methods

Recruitment

Recruitment was done through snowball sampling, using social media platforms (eg, Facebook). A Google Form survey was developed containing the informed consent and the questions and questionnaires needed to evaluate the women's eligibility for participating in the study. Specifically, eligible women complied with the following inclusion criteria: they (1) had physiological pregnancy, (2) were aged ≥18 years, (3) were between the 12th and 30th week of gestation, and (4) had subclinical depression symptoms (Patient Health Questionnaire-9 [PHQ-9] score<15) [34]. By contrast, women were excluded when (1) presenting a history of past or current mental disorders; (2) exhibiting clinically significant psychological symptoms (ie, depression symptoms: PHQ-9≥15) and suicidal ideation (PHQ-9 item 9); (3) having an obstetrically at-risk pregnancy; (4) presenting medical conditions, pregnancy-related and otherwise; (5) experiencing an artificially induced pregnancy. A total of 15 women had filled in the web-based questionnaire; all were deemed eligible, and thus none reported any of the exclusion criteria. Following randomization in either the guided or unguided group, 4 (27%) women dropped out (Figure 1) before starting the intervention. As such, the final sample is composed of 11 (73%) women.

XSL•FO
RenderX

Figure 1. Study structure and study adherence flowchart. *Indicates optional modules and # indicates that 1 participant had stopped interacting with the guide but had continued viewing the Moodle content. BADS-SF: Behavioral Activation for Depression Scale–Short Form; EPDS: Edinburgh Postnatal Depression Scale; EROS: Environmental Reward Observation Scale; GAD-7: Generalized Anxiety Disorder-7; PHQ-9: Patient Health Questionnaire-9; PSS: Perceived Stress Scale; T: time point; UE: user engagement; UX: user experience.



Procedure and Study Structure

This study is structured in 4 main phases (Figure 1). Phase 1 corresponds to time point 0 (T0), which encompasses the recruitment, enrolment (including the assessment of inclusion and exclusion criteria), and randomization processes. At this time, anamnestic information (reported in the *Recruitment* section) was collected, and standardized questionnaires were administered measuring depression and anxiety symptoms, perceived stress, current activity level, and perceived environmental reward. Phases 2 and 3 of the study correspond to T1-T6 and T7-T9, respectively. During these phases, the baseline questionnaires, along with one questionnaire assessing user engagement (UE) and another assessing user experience (UX; explained in the *Measurement Tools* section) were administered. In addition, the UX questionnaire was also administered during T2, T4, and T8. Questionnaires assessing the level of activity and related reward (explained in the *Measurement Tools* section) were administered each week, from T0 to T9. The final phase 4 involved conducting semistructured interviews that were created ad hoc, and participants who had participated up to at least T6 were interviewed. The semistructured interviews were conducted to qualitatively evaluate the women's experience with the intervention and gather further feedback necessary for refining the intervention. All questionnaires administered between T1 and T9 were created using Google Forms. Links to access these questionnaires were made available within the intervention platform and were accessible to both groups in the same manner.

The participants were informed that they could leave the study at any moment without having to provide an explanation and without incurring in any penalty. Each was assigned an alphanumeric code to ensure confidentiality.

Randomization

Alphanumeric codes were generated and allocated to each participating woman to guarantee confidentiality throughout the study. The process of creating these codes was carried out using Microsoft Excel and further randomized through a Google software [35], ensuring the unbiased assignment of participants to either the guided or unguided group before the commencement of recruitment. The results of random code assignment were kept aside, and women were only provided with their designated code once their eligibility for the study had been established. Only when eligibility was confirmed, participants were given specific information on the intervention they had to follow. The group assignment was single blinded.

The Intervention's Content

This internet-based self-help intervention originates from the BATD-R [31] protocol. Originally, this protocol consisted of 10 sessions, which included 5 main sessions and 5 additional sessions aimed at reviewing and maintaining the benefits achieved; nevertheless, it is possible to reduce the number of sessions to 5 or even expand beyond 12 sessions. However, past studies advise against exceeding 10 weeks of intervention, reporting that BATD-R interventions of up to 6 or 8 weeks allow for a significant reduction in depression symptoms [36,37]. BATD-R can be self-administered or administered by both trained and untrained staff, further emphasizing its adaptability and flexibility. The protocol begins with a psychoeducation phase focused on understanding the characteristics of depression symptoms. Furthermore, it includes

XSL•FO
RenderX

homework assignments to be completed between sessions, which form the core content of the 5 main sessions. These assignments involve filling out 5 forms (refer to the original protocol by Lejuez et al [31]). The first form, called the "daily monitoring form," should be completed throughout the intervention and requires participants to continuously evaluate their daily activities, in terms of behavioral patterns, the pleasantness of activities, and the importance of each daily action. Subsequently, the person is required to identify through a second form, called "life areas, values, and activity inventory," what their values are within important areas of life (ie, important relationships; pleasurable activities and hobbies; work and study; mind, body, and spirituality; and daily responsibilities). This would then allow the person to identify (form 3, called "activity selection and ranking") and plan daily activities (form 4, called "action plan") that help them live in accordance with their values in the mentioned life areas. In this regard, it is worth noting that although the BATD-R does not include a complete functional analysis due to the brevity of the treatment approach [38], Lejuez et al [31] stressed that several components of treatment fit into a functional analytical framework. This is most noticeable when choosing activities that are closely related to values, given the dual goals of identifying the factors that maintain or reinforce depressive behavior (both positive and negative reinforcement) and the positive reinforcers that could support or strengthen healthy behavioral patterns. Along with the critical assessment of dysfunctional behavioral patterns, the BATD-R protocol includes a final form called "contracts," which focuses on strategies to request social support. To create these "contracts," the person is required to identify (1) an activity to perform, (2) up to 3 support persons who may be able to assist or support them, and (3) how and when each person can specifically provide said support. This activity allows the patient to identify their needs and provides a specific plan on how to seek the help they need by making concrete requests for obtaining assistance and social support to reinforce adaptive behaviors.

In this study, the intervention closely follows the structure of the original protocol but is adapted to fit a digital format. The intervention was divided into 6 weekly sessions or modules, which include the core content of the intervention. In addition, 3 optional "bust" sessions were included to reinforce and consolidate the information from previous weeks. The entire intervention spanned 9 weeks, with participants completing weekly homework assignments between sessions. While the intervention content and homework remained unchanged, it was adapted for digital delivery using text, video, and images. Information related to depression symptoms was contextualized for the pregnancy period to cater to the specific needs of the target population.

The Platforms

Overview

In this first evaluation of the intervention, the Moodle e-learning platform (Moodle 3.11; 2021) was used as the delivery method. This platform was accessible via both the web and the Moodle app on smartphones. Both the web and app versions of Moodle included a chat feature, which was exclusively used by the guided group to interact with their assigned guides. Specifically, in the guided group, guide-woman dyads were created, and they interacted once a week for the homework revision. For the unguided group, homework revision was facilitated through a written self-guide available within the Moodle platform. The forms representing the homework were structured on Google Docs, with links to access them made available within the Moodle platform so that women could directly access them both on the web and the smartphone app.

Both intervention groups were presented with the intervention as an e-learning Moodle course; the sole difference in the intervention's content between the groups lay in the reference to the guides. The intervention was structured into modules (6 core modules and 3 optional modules), and each module could be consulted only after completing the previous one. The content of each module was delivered using illustrative videos and images, complemented by brief text information. After viewing each section within a module, participants were presented with a brief quiz comprising 3 true-or-false questions related to the content they had just reviewed. These quizzes were incorporated with the intention of fostering UE. They encouraged participants to actively engage with the material rather than to passively view it. In addition, the quizzes served as a means of assessing participants' comprehension of the content. On the basis of the accuracy of their responses, participants received reinforcing or motivating feedback after completing each quiz.

The Guides

In the guided group, specific guide-woman dyads were randomly created. All the guides were recognized psychologists who had been trained to become psychotherapists. They underwent comprehensive training, which included the provision of detailed written information and a 2-hour in-person meeting. This training covered the intervention's content and structure and their role as guides. The training aimed to ensure that all the guides had a consistent understanding of the intervention and its content, thus maintaining uniformity across the interactions between the guides and participants. The guides adhered to a partially defined conversational protocol, which consisted of fixed messages and information to be delivered, as well as "free" parts where they had the freedom to phrase sentences as they saw fit. This flexibility in the "free" parts allowed guides to respond adaptively to women's answers and feedback, particularly during the homework revision part of the intervention. Furthermore, it enabled guides to support participant compliance and adherence based on their perceived motivation levels. Supervision for the guides was provided by the first author (EM) of the study, who oversaw the technical aspects of the intervention. In addition, an expert psychotherapist, the third author (SS), provided supervision for clinical matters.

To ensure a consistent participant experience, all the guides were assigned the same name, "Joy." This uniformity aimed to minimize any potential biases that could arise from variations in the perception of the different guides. Moreover, guides were not provided with any information about the women they were assigned to, ensuring privacy and confidentiality for the participants.

XSL•FO
RenderX

Measurement Tools

PHQ-9 Tool

The PHQ-9 [34] is a unidimensional self-report tool that is widely used in the Italian context [39]. It assesses the severity of depression symptoms during the previous 2 weeks, based on the diagnostic criteria of the *Diagnostic and Statistical Manual of Mental Disorders* (*Fourth Edition*; *DSM-IV*) [40]. It consists of 9 items measured on a 4-point Likert scale (0="not at all"; 3="almost every day"). Item 9 assesses suicidal ideation. A score of ≤9 indicates mild or no symptoms of depression, and a score between 10 and 14 indicates moderate symptoms, while a score of ≥15 indicates severe symptoms of depression. The instrument shows excellent internal consistency at α=.92 [41].

Edinburgh Postnatal Depression Scale

Edinburgh Postnatal Depression Scale [42] is a unidimensional self-report tool, validated in Italy [43], which assesses the severity of depression symptoms during the previous week. Albeit developed to assess depression symptoms during the postpartum period, it is often used throughout the perinatal period. It consists of 10 items measured on a 4-point Likert scale (0="no, not at all"; 3="yes, always"). The instrument shows good internal consistency at α=.79 [43].

Generalized Anxiety Disorder-7

Generalized Anxiety Disorder-7 [44] is a unidimensional self-report tool that assesses the severity of anxiety symptoms during the previous 2 weeks. It consists of 7 items measured on a 4-point Likert scale (0="never"; 3="almost every day") and shows good psychometric indexes in the Italian context as well [41]. The instrument shows excellent internal consistency at α=.92 [41].

Perceived Stress Scale

Perceived Stress Scale [45] is a unidimensional self-report tool, validated also in Italy [46], which assesses the severity of stress symptoms in the previous month. It consists of 10 items measured on a 4-point Likert scale (0="never"; 3="quite often"). The instrument shows good internal consistency at α=.74 [46].

Behavioral Activation for Depression Scale—Short Form

Behavioral Activation for Depression Scale–Short Form [47] is a self-report tool designed to measure changes in avoidance and activation during BA interventions for depression during the previous week. It consists of 9 items measured on a 7-point Likert scale (0="not at all"; 6="completely"). The scale provides 2 scores, the first score referring to the level of BA (5 items) and the second one to the level of behavioral avoidance (5 items). Manos et al [47], the authors of the tool, advise considering the total score instead of the subscales. This questionnaire has not been translated into Italian and was therefore translated through the back translation procedure. Example items are "I am content with the amount and types of things I did" (item 2) and "Most of what I did was to escape from or avoid something unpleasant" (item 5). The instrument shows good internal consistency (total scale α=.82) [47].

Environmental Reward Observation Scale

Environmental Reward Observation Scale [48] is a unidimensional self-report tool designed to measure the level of environmental reward perceived in recent months. It consists of 10 items rated on a 4-point Likert scale (0="strongly disagree"; 4="strongly agree"). This questionnaire has not been translated into Italian and was therefore translated through the back translation procedure. Example items are "It is easy for me to find enjoyment in my life" (item 4) and "I wish that I could find more hobbies that would bring me a sense of pleasure" (item 7). The instrument shows good internal consistency (α=.87) [48].

UX Measure

Mobile Application Rating Scale (MARS) [49] is a self-report tool consisting of 23 items scored on a 5-point Likert scale (1="poor"; 5="excellent"), which assesses the quality of the app and its features (ie, the Moodle app) on 4 dimensions of objective quality: engagement (5 items), functionality (4 items), aesthetics (3 items), and information (7 items); a final scale assesses the subjective quality (4 items). The average of the scores of the 4 dimensions of objective quality provides the total scale score. The questionnaire also contains an "application-specific" section (6 items) to assess the potential impact of a particular app on domains such as users' knowledge and intentions. The total and subscale scores of the MARS have high internal consistency coefficients (α=.90 and α=.80-.89, respectively). The scale has been validated in the Italian context [49]. For this study, only the subscales related to "information," "subjective app quality," and "app-specific" sections were considered, totaling to 17 items.

Together with the MARS items, only at T6 and T9, women were also asked to prove their overall subjective opinion on the platform ("Please write below your personal opinion with respect to your experience [pros and cons] while using the platform").

UE Measure

UE Scale–Short form [25] is a short self-report tool designed to assess UE with a digital solution. It consists of 12 items based on a 5-point Likert scale (1="strongly disagree"; 5="strongly agree"). The questionnaire consists of four factors: (1) focused attention, which indicates the feeling of being immersed in the interaction; (2) perceived usability, which is the negative effect experienced due to the interaction and the effort expended; (3) aesthetic attractiveness, which represents the graphical and visual appeal concerning a digital solution; and (4) the reinforcement (reward) factor, which regards the perceived involvement and enjoyment with the digital solution. This questionnaire was not translated into Italian and was therefore translated through the back translation procedure. The 4 scales have good internal reliability, as follows: focused attention, ω=0.75; perceived usability, ω=0.70; aesthetic attractiveness, ω=0.88; and reinforcement, ω=0.79 [25].

Semistructured Interview

The semistructured interviews were conducted by the first author (EM) and featured 15 main questions developed specifically for the study, 3 (20%) of which were asked only to the guided

XSL•FO
RenderX

group. This interview was conducted approximately 10 days after each woman had finished the intervention. It lasted between 15 and 20 minutes, and following the woman's consent, it was audio recorded to allow for its transcription and evaluation. For both the guided and unguided groups, the interviews investigated women's personal experience (7 questions) with the intervention and the experience (5 questions) specifically related to the use of the platform. In the guided group, women's experience with their guide and the overall chat interactions were investigated.

Data Analysis

Statistical analyses were performed using RStudio (R Foundation for Statistical Computing) [50]. Descriptive data for both categorical (n, %) and continuous (mean and SD) variables were analyzed separately for the guided and unguided groups, considering the different time points. Given the preliminary nature of the study and the small sample size, no further analyses were performed. The interviews were individually and qualitatively analyzed through thematic analysis, following the predefined semistructured interview's 3 broader themes. Thematic analysis was conducted following a modified version of the guidelines proposed by Braun and Clarke [51], which has already been used in other co-design studies [52].

Ethical Considerations

The study was conducted in compliance with the ethical guidelines of the Declaration of Helsinki [53] and the European Union law for data protection (EU General Data Protection Regulation 679/2016). The study was approved by the Ethical Committee of the Psychology Department of the University of Padova (number 4820/2022).

Results

Descriptive Information and Adherence

A total of 11 women participated in the study; 6 (55%) were part of the guided group and 5 (45%) were part of the unguided group. Descriptive information is reported in Table 1 separately for the 2 groups.

Table 1. Descriptive information (n=11).

	Guided group (n=6)	Unguided group (n=5)
Age (years), mean (SD)	32.17 (4.36)	31 (4.95)
Gestation week, mean (SD)	21.17 (5.95)	19.83 (5.75)
Living area, n (%)		
North Italy	4 (67)	4 (80)
Central Italy	2 (33)	1 (20)
South Italy	0 (0)	0 (0)
Education, n (%)		
<High-school diploma	1 (17)	0 (0)
High-school diploma	1 (17)	2 (40)
Bachelor degree	1 (17)	0 (0)
Master degree	2 (33)	2 (40)
Specialization (eg, PhD)	0 (0)	1 (20)
Marital status, n (%)		
Single	0 (0)	0 (0)
Cohabitant	3 (50)	2 (40)
Married	3 (50)	3 (60)
Past abortion, n (%)		
Yes	2 (33)	1 (20)
No	4 (67)	4 (80)
Women's occupation, n (%)		
Unemployed	1 (17)	0 (0)
Student	1 (17)	0 (0)
Student and freelance worker	1 (17)	0 (0)
Employee	3 (50)	3 (60)
Student and employee	0 (0)	1 (20)
Researcher	0 (0)	1 (20)

XSL•FO
RenderX

The descriptive statistics pertaining to psychosocial variables assessed at T0, T6, and T9 for both groups are presented in Table 2.

Table 2. Descriptive statistics at time point 0 (T0), T6, and T9 (n=11).

	Scores of the guided group (n=6)			Scores of the unguided group (n=5)	
	T0 (n=6), mean (SD)	T6 (n=3), mean (SD)	T9 (n=1)	T0 (n=5), mean (SD)	T6 (n=1)
PHQ-9[a]	6.67 (2.66)	3.33 (2.08)	3	3.4 (2.79)	1
EPDS[b]	15 (2.37)	11 (3.61)	6	10.6 (3.21)	9
GAD-7[c]	7 (1.41)	5 (2)	3	4.0 (1)	3
PSS[d]	20.50 (2.07)	18.67 (2.08)	17	18.0 (2.3)	17
BADS-SF[e]	23.17 (8.18)	16.67 (1.15)	18	19 (2.92)	27
EROS[f]	28.17 (4.17)	30.67 (2.52)	34	32 (3.39)	36

[a]PHQ-9: Patient Health Questionnaire-9.

[b]EPDS: Edinburgh Postnatal Depression Scale.

[c]GAD-7: Generalized Anxiety Disorder–7.

[d]PSS: Perceived Stress Scale.

[e]BADS-SF: Behavioral Activation for Depression Scale–Short Form.

[f]EROS: Environmental Reward Observation Scale.

Regarding dropout rates, it was higher in the unguided group (n=5), with most participants (n=4, 80%) dropping out after completing the second module. In contrast, the dropout pattern in the guided group (n=6) was more gradual. Overall, 1 (9%) participant had dropped out because of health reasons and 3 (27%) because of the amount of time and effort (particularly related to the homework) required by the intervention, while 3 (27%) did not provide a reason for dropping out. Accordingly, 3 (50%) participants from the guided group reached T6, thereby completing the 6 core modules in Moodle, while only 1 (20%) in the unguided group reached T6. However, among the 3 participants from the guided group who reached T6, 1 (33%) ceased interactions with her guide after T3 but continued viewing the material on Moodle up to T6; the other 2 (67%) continued until T9. One of them stopped viewing the Moodle content at T8 but continued with the chat interactions with the guide.

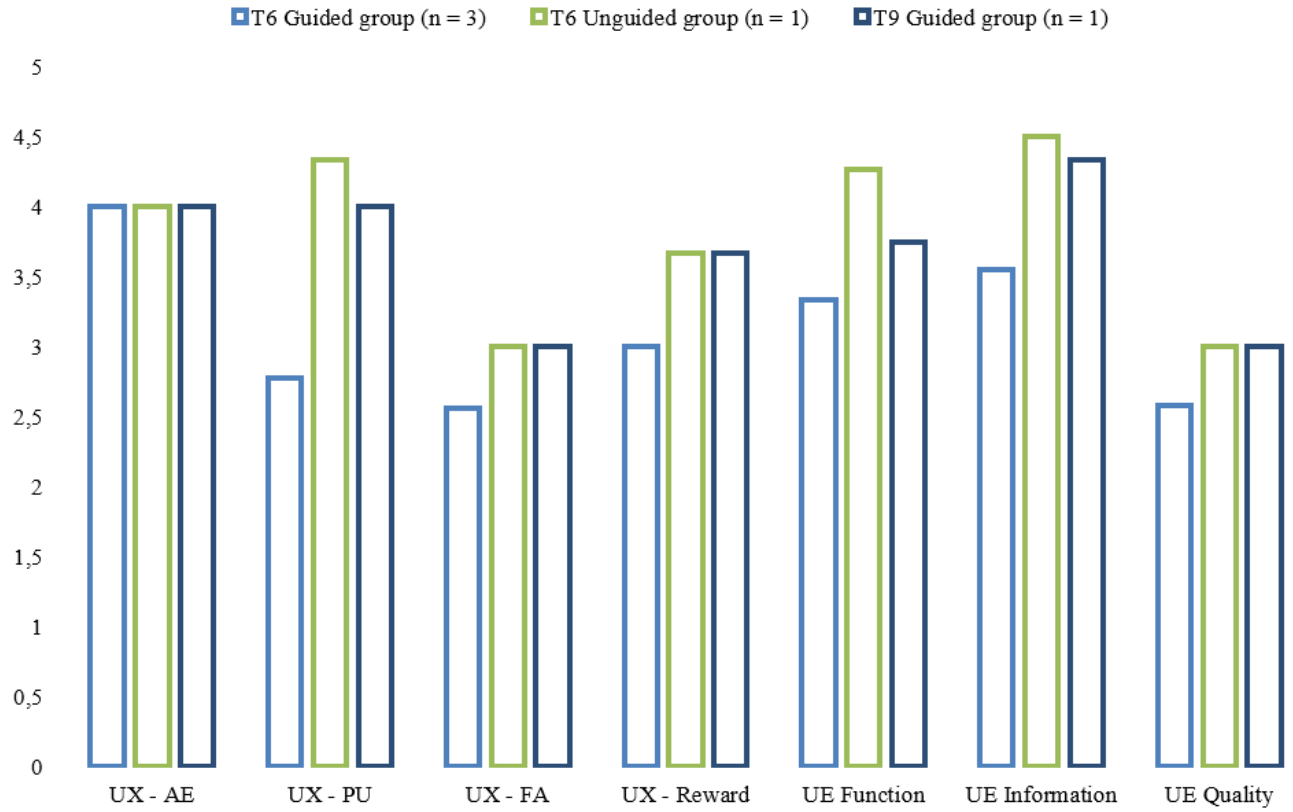
Regarding homework completion, it is not possible to quantify adherence specifically, as, for instance, among the participants that had completed at least until T6 some (2/4, 50%) decided to handwrite the homework, instead of using Google Docs because of the low usability of the latter (ie, too many steps to go from Moodle to Google docs, which were also not well perceived and difficult to use within the smartphone). Moreover, it was considered cumbersome to write within the Google Docs.

UX and UE Measures

Descriptive statistics for UX and UE assessed at T6 and T9 are illustrated in Figure 2. At T6, when the participants were asked whether they had used primarily the app or web version of Moodle, 2 (67%) participants of the guided group and the only 1 participant of the unguided group reported using the app version, while 1 (33%) of the participants of the guided group used primarily the web version because of difficulties with using the app.

XSL•FO
RenderX

Figure 2. User experience (UX) and user engagement (UE) at time point 6 (T6) and T9. All scales' response range was from 1 to 5. AE: aesthetic appearance; FA: focused attention; PU: perceived usability; quality: subjective quality.



Participants' subjective opinion on the platform, assessed through one open question, is reported in Table 3.

Table 3. Answers to the open question "Please, write below your personal opinion with respect to your experience (pros and cons) using the platforms."

Time points and group	Participants' feedback
Time 6	
Unguided group, n=1	• "I am in my second pregnancy so I have been using the platform with a one-and-a-half-year-old [taking care of them] and I must say that being able to subdivide the time has been helpful even on an organizational level. If one fills out the daily forms from day to day it is not challenging, I must say that filling them out [the daily forms] from your cell phone though is quite inconvenient because the app [Google Docs] that opens the forms and allows you to fill them out does not always work well, so I then preferred to print them out and fill them out by hand."
Guided group, n=3	• "I think it is still very cumbersome as a platform [Moodle], not easy to use and the interaction with the Guide [was] too dry. Some things need to be revised." • "Using the platform on a practical level was quite intuitive and easy. The project itself involves a lot of effort and concentration, but it helps to feel a greater physical and psychological well-being." • "It was intuitive and fast."
Time 9	
Guided group, n=1	• "My experience in using the platform has been positive."

Semistructured Interviews Results

A total of 3 participants agreed to participate in the semistructured interview; 2 (67%) were from the guided group and had continued the intervention till T9, while 1 (33%) was from the unguided group and had participated till T6.

During the semistructured interview, all 3 (100%) participants reported that they had come to know of the study through a friend. As motivation for participating, all 3 reported "*curiosity*" as the main reason because they did not have specific expectations before starting.

Overall, participants expressed that the intervention helped them find a moment for themselves and provided them with a method or strategy to change their perspective on how they viewed and performed their daily activities. Furthermore, they noted the positive effects of engaging in more rewarding activities. All 3 (100%) participants emphasized that the intervention supported their overall well-being rather than reduced negative feelings

XSL•FO
RenderX

per se. However, they also mentioned that the effort required by the intervention, particularly in terms of time and dedication, was substantial, especially regarding the homework assignments. While they appreciated the meaningful content of the homework, they found it burdensome to complete. It is important to note that part of the difficulty with homework completion was related to the reduced fluidity and usability of Google Docs on smartphones. A more thorough explanation of the findings that emerged from the semistructured interviews with verbatim examples of participants' answers has been discussed as follows, and the specific themes and subthemes that emerged are reported in Table 4.

XSL•FO
RenderX

Table 4. Semistructured interview (n=3): themes, subthemes, and examples.

Theme and subtheme	Example quotes
Personal experience with the intervention	
Self-observation and activities evaluation	• "[...]you realize things that by not doing it [the intervention] you wouldn't have realized you could have done, how you could have also handled the pregnancy period better" [Unguided group participant] • "It definitely helped me understand what are [...] let's say, [which is] the focus, where to aim to get better. I discovered some things that I had set aside[...] it helped me to not be so focused on negative things, which is kind of my problem, but more to identify something positive to do day-by-day, to be able to accomplish little goals that [might not be important for others] but for me at that moment they were important" [Guided group participant] • "[...]most definitely [I appreciated] focusing my attention on some positive activities that I had somewhat set aside and forgotten[...] when I focused on those saddest moments and I [then] realized that they were not most of my days as one thinks when one is in the sad mood. Instead, I saw that most of my days were good moments" [Guided group participant]
Effort	• "As the weeks went on, maybe even as the pregnancy progressed[...] I found it more 'burdensome.' The fact of filling out the daily forms [daily monitoring form] every day[...]. as time went on, it was challenging, in the sense that one has to stop and really take [their] time and be consistent, when as the pregnancy progresses maybe other thoughts take over and you can't quite be that consistent all the time[...]" [Guided group participant] • "[...]the part, let's say the most obnoxious, difficult, whatever we want to call it, is definitely the material to fill out during the last weeks. I have to tell you the truth, I didn't even finish them because I didn't have time[...]" [Guided group participant] • "In my experience[...]. I felt the fatigue more, maybe, here. Let's say that at the end of the six (sixth module), for myself, I felt that the intervention was -in quotes- "finished." [Guided group participant] • "[In reference to the intervention length] it probably depends on what stage of the pregnancy one is[...] I started it toward the end of my pregnancy, it would probably be better to start it before[...]"[Unguided group participant]
Learning the "method" and its application in the future	• "[...]It might be a good method at other times in life when one may experience difficulties" [Guided group participant] • "[...]mentally, I got into this mode of planning something nice to do, to be able to have that weekly commitment that I like, to ask somebody to do it with me. Maybe small things, but I'm sure it helped me for the future as well." [Guided group participant]
User experience and user engagement	
Managing issues and learnability	• "[...]when I had to write [for the homework], there was the transition from the daily monitoring [form] rather than very often I had trouble writing things down[...] I mean, I had to print them out [the forms] basically, if not I couldn't record [write] them [down]." [Unguided group participant] • "[A]t first maybe you kind of have to learn the mode of, yeah, that you get the materials out[...] It wasn't easy to tell when it saved what you had done, because it was a little dubious, sometimes[...] it was easy to use once you had gotten the hang of it, going into the week, looking at the material..." [Guided group participant]
Multimedia material	• "[...]the videos etc. were very effective in passing the message, both in terms of explanations and content, and the images, they were[...] they caught the attention[...]" [Guided group participant] • "[T]he videos[...] they are very clear, well done, they are cute" [Unguided group participant]
App interventions in the future	• "[...]with the current use of the phone and the computer in general, in my opinion an app is useful" [Unguided group participant] • "[...]the app that you have on your phone is the most useful thing. You consult it wherever you want and whenever you want" [Guided group participant]
The experience with the guide	

XSL•FO
RenderX

Theme and subtheme	Example quotes
Support	• "[...]the interaction itself was effective in the sense that it explained things to me when I had doubts, that it directed me, maybe, when I didn't quite understand the task, it directed me well[...]" • "[...]giving me suggestions, helping me even on how not to give up-because maybe there were also harder moments-thus also giving me alternatives and suggestions to see the path in a different way, make it lighter[...] I had a great time"
Personalization of the conversation	• "[...]the redundancy of the messages[...] sometimes it almost felt like a copy and paste of the messages and not an actual interaction[...] I found it a little depersonalized[...] I would have found it more enjoyable if it was personalized[...]" • "I saw that the person who was on the other side—I don't like to say 'Joy the guide!'—I still found her to be a person who each time, with respect to my mood, to how my week had gone, has put herself into my shoes, into my being, into my experience[...]"

Participant's Personal Experience With the Overall Intervention

Coherent with the *self-observation and activities evaluation* subtheme reported in Table 4, participants reported a positive personal experience as they seemed aligned with the content of the intervention, thus learning to appreciate the value of self-observation (regarding behavior and emotions) and how this can influence how they feel. For instance, the participant from the unguided group reported as follows:

> *[...]it [the intervention] was helpful, because it may not seem like it but by writing down daily what you do... you notice things that maybe normally you wouldn't notice, or [discover] free time that you can spare, which maybe you didn't even think about...*

Coherently, a participant from the guided group affirmed as follows:

> *[...]in relation to well-being, it definitely helped me understand[...] where to aim to get better... being at home during pregnancy, I rediscovered some hobbies that helped me... it helped me to not be so focused on negative things[...] to identify something positive to do day-by-day, to be able to accomplish little goals that [might not be important for others] but for me at that moment were important.*"

Participants seem to have internalized the "message" that the intervention wanted to transmit. When asked whether they believed what they had done during the intervention could be useful to them in the future, participants reported already having integrated what they have learned into their daily life, for instance, by starting "[...]to keep a journal[...]" (a participant from the unguided group), even after pregnancy. Notably, a participant from the guided group reported as follows:

> *[E]ven in conditions of not pregnancy... it might be a good method at other times in life when one may experience difficulties.*

Referring to the subtheme of *learning the "method" and its application in the future*, this highlights that the intervention content was able to provide the participants with a broader method useful to support their psychological adjustment and well-being both during pregnancy and in the future:

> *[...]it's like, mentally, I got into this mode of planning something nice to do, to be able to have that weekly commitment that I like, to ask somebody to do it with me. [These] may be small things, but I'm sure it will help me in the future as well.* [Guided group participant]

In this regard, and referring to the *effort subtheme*, participants reported that 6 weeks of intervention were enough to this end, while the subsequent optional 3 weeks were perceived as a bit redundant and excessive. In addition, they stressed the effort required by the intervention overall, affirming, for instance, as follows:

> *[A]s the weeks went on[...] I found it more burdensome[...] filling out the daily form chart [daily monitoring form] every day with the activities perhaps, as time went on, was challenging, in the sense that you have to stop and really take [your] time and be consistent, when as the pregnancy progresses maybe other thoughts take over and you can't quite be that consistent all the time[...]* [Guided group participant]

On a similar note, another participant reported as follows:

> *[...]the part, let's say, the most obnoxious, difficult... is definitely the material to fill out during the last weeks. I have to tell the truth; I didn't even finish them because I didn't have time[...]* [Guided group participant]

Finally, coherent with the emerging subthemes, a further comment made by a participant from the guided group ought to be reported; she stressed that women would need to be already motivated to be able to appreciate the intervention. Indeed, she mentioned a key point, which is the need to have an adequate capacity for insight, as it might otherwise be difficult to autonomously notice the maladaptive behavioral patterns and switch to more positive ones. In this regard, the participant reported as follows:

> *[...]they [those following this intervention] must be people[...] who are capable of introspection... I imagine people who don't have so much of a way of knowing themselves[...] it's not so easy to start on such a path if you haven't done some work on yourself first[...] also, just to have self-knowledge and say*

XSL•FO
RenderX

> *"what are my weak points?" and say "where can I go with that?"*

UX and UE Themes

Regarding UX and UE, the *managing issues and learnability* subtheme was quite prominent. All 3 (100%) women agreed that the Moodle app was not so easy to use overall, as some participants' approach has been that of learning how to work around what was not working to be able to continue the intervention. For instance, within an otherwise positive experience, a participant reported as follows:

> *...maybe the only thing that I would change, which is not really of the intervention though, [regards] more the use of the platform, is just that[...] I had trouble writing things down [for the homework...] even [the speed of the] connection when it makes you log back in to do the quiz, the page has to reload[...].* [Unguided group participant]

Similarly, a participant from the guided group reported as follows:

[Although she was a] *geeky chick* [as regards to the use of technology], *it was not easy* [to use the Moodle platform]. *It took me a while to find the material[...] I had even emailed you* [the researcher conducting the interview]... *I couldn't really understand how it worked, where they* [the materials] *were[...] I always used it from the smartphone, only a couple of times I used it from my computer[...] maybe you kind of have to learn first* [how to use the platform][...] *then I had that glitch with the quiz, and that one I found a little obnoxious, because I thought I had done the quiz[...]. it had not saved nothing. It wasn't easy to tell when it saved what you had done, because it was a little dubious, you know, sometimes[...] it was easier to use once you had gotten the hang of it[...].*

Altogether, this stresses the importance of the platform's simplicity, ease of use, and related learnability regarding UX and UE. However, it should be noted, referring to the *multimedia material* subtheme, that the aesthetic of the material present, and particularly the videos and images, were much appreciated and perceived as informative.

Furthermore, coherent with the *app interventions in the future* subtheme, all 3 (100%) participants reported that they did believe that a smartphone app, being readily available, could be a valuable tool to administer this sort of intervention saying, for instance, that "...with the current use of the phone and the computer in general...an app is useful" (the participant from the unguided group) as one can "...consult it wherever and whenever..." (a participant from the guided group). Indeed, albeit reporting difficulties with the platform, all 3 (100%) participants had already autonomously recommended the intervention to fellow pregnant women. However, the moving force, coherent with what is reported earlier, was the intervention content:

[Although deployed through a portable tool, interventions] *help in times of transition or change. This tool* [the present digital intervention] *helps because it focuses on pleasant activities, and in such a long waiting time* [the pregnancy], *with the struggles to organize exams, visits, what's going to happen tomorrow[...] it helps you a little bit to[...] focus on simpler thing, that then in itself just help you every day. So yes, I would recommend it for that* [Guided group participant]

The Experience With the Guide

Regarding the role of the guide and the guided group's overall experience with the chat interactions, the *support* subtheme has emerged, as both participants reported the positive value of having this sort of support, whether it be practical or affective. However, intersecting with this support subtheme, the *personalization of the conversation* subtheme seemed to have weighted on participant perception of the quality of the support perceived. In particular, it seems plausible to hypothesize that as participants knew that the guide was an actual person, the way of talking of the guide in the free sections of their protocol guided the participants' perception of their capacity for empathy and perception of getting in tune with them. Indeed, while 1 (33%) of the 3 participants reported the intervention to be not personalized enough, 1 (33%) reported almost the opposite. More specifically, a participant reported as follows:

> *[...]the interaction itself was effective in the sense that it explained things to me when I had doubts, that it directed me, maybe, when I didn't quite understand the task, it directed me well. Um, the part that I definitely didn't like was the redundancy of the messages because sometimes it almost felt like a copy and paste of the messages and not an actual interaction[...] I found it a little depersonalized[...] I would have found it more enjoyable if it was more personalized.*

The other, instead, reported expecting from the beginning a set of predetermined questions from the guide, particularly as some questions were indeed repeated each week; however, the participant reported about the guide as follows:

> *[S]aw that the person who was on the other side, I don't like to say "Joy the guide"- I found them to be a person who each time, with respect to my mood, to how my week had gone, has put herself into my shoes, into my being, into my experience, giving me suggestions, helping me even on how not to give up-because maybe there were also harder moments-thus also giving me alternatives and suggestions to see the path in a different way, to make it lighter. So, I had a great time!*

Such difference in the perception of the guide and of their helpfulness seems even more plausible when considering that a participant (who did not agree to the semistructured interview) had finished viewing the modules on Moodle until T6 but did not continue with interactions with the guide after T3.

Discussion

Principal Findings

This study aimed to investigate the feasibility of the BATD-R protocol [31] that was structured as an internet-based self-help intervention and deployed to pregnant women with subclinical depression symptoms, while further comparing its guided versus

XSL•FO
RenderX

unguided versions. Such evaluations had the associated purpose of collecting the feedback needed to refine the intervention, thereby allowing its co-design and adaptation; moreover, it represents the first instance where the replication of the in-person BATD-R into a digital format has been empirically measured and evaluated, yielding valuable insights and shortcomings useful for future research.

Overall, results showed that this first version of the digital BA intervention, JuNEX, as a "replica" of the original BATD-R protocol [31], is not feasible to be implemented in a digital e-learning format. Specifically, data have highlighted the need to lighten the intervention and reduce the effort required for homework. Despite the perceived excessive effort, participants appreciated the content and purpose of the intervention, which allowed them to "take a moment" for themselves and understand "where to aim to get better." Indeed, albeit the intervention protocol originated from a behavioral framework, its focus on the person's everyday activities, the evaluation of one's own experiences, and how these are linked to how behaviors and emotions mutually influence each other configure in line with third-generation cognitive behavioral therapies [54]. These therapies prioritize holistic enhancement of psychological and behavioral processes related to health and well-being [55], emphasizing adaptive coping methods and increasing experiential and contextual awareness [55,56]. Such an approach is particularly relevant for nonclinical populations (which are more diverse than clinical populations), as they allow for a more transversal relevance and application of the coping methods promoted, thus making the intervention especially valuable in preventive terms. Given that this study focused on women with subclinical depression symptoms, interventions emphasizing the awareness of psychosocial functioning and the interplay between emotions and behaviors may be more beneficial than targeting specific, limited areas of functioning and distress. This broader approach could enhance women's capacity for adjustment and could be applicable to difficult situations beyond pregnancy-related challenges.

Such explanations serve to emphasize that consistent with women's feedback, the intents of the intervention per se, and thus its content as well as homework purposes, ought to be maintained as they were found pertinent to the pregnancy situation and were appreciated by women; however, data also stress the need to structure the intervention so that it can be deployed with greater ease and the need to shorten it to maximum of 6 weeks. In this regard, the data do point to satisfactory usability and UX as pivotal aspects of the intervention feasibility, which was not provided either by the Moodle platforms or the Google Docs used. These platforms were used in line with the preliminary and exploratory nature of this study, allowing the first structuring of the BATD-R protocol in a digital setting in a time- and cost-efficient manner; this has favored the co-design of the future development of the intervention by developing a more advanced and refined application after having collected some initial pregnant women's feedback.

When evaluating the usability of a digital solution, 5 main aspects are to be considered as follows [57]: (1) the simplicity of use experienced by users when learning how to use a digital tool, (2) the number of mistakes they make to do a certain action correctly, (3) the effectiveness with which users interact with a digital tool, (4) the perceived satisfaction with the UX, and (5) the memorability of how to use a digital tool after having been exposed to it. In addition to this, and particularly linked to both the effectiveness and perceived satisfaction just mentioned, is the more specific UE, thus linked to the subjective experience of the user and subsuming affective, cognitive, and behavioral components [25]. Considering the UX and UE evaluated in this study, data suggest that usability and UE were mediocre, yet not scarce. Thus, there was some appreciation for the tools used by the participants who had completed the interventions; however, the simultaneous high dropout and overall low recruitment of participants strengthen the idea that the more positive feedback given is somewhat linked to the a priori internal motivation of the participants in following the intervention. Participants themselves have indeed underlined that to follow the intervention as it is, to be able to bear the effort required by it, and to work around the limits of the platform used, women would need to be highly motivated on their own.

Coherent with the importance of the users' motivation to properly follow such interventions, data do suggest that the guided group showed greater adherence and were overall more willing to finish the intervention compared to the unguided group. A part of the guides' job was indeed to motivate women to favor compliance and adherence, and as women knew that they were interacting with a psychologist, this can be thought to have further increased their motivation. However, the practical challenges of having to set a date and time for the interactions are the limiting factors and so are the individual differences related to the guides' different writing modalities. In this regard, a step further in this direction might be the design and inclusion of a conversational agent to provide guidance and support during the digital intervention. Conversational agents can greatly favor the personalization of the user-system interactions while fostering scalability by requiring a much-reduced workforce for the intervention administration [58]. Existing literature has highlighted that conversational agents might be valuable tools to foster intervention adherence by favoring engagement and involvement, thereby supporting the overall UX [58,59] and allowing for a more immersive experience. Using a conversational agent is expected to further reduce time constraints for both patients and clinical professionals while also reducing health care costs in the long run. Within the perinatal context, future studies should thus evaluate the potentiality of including a conversational agent within digital interventions to ultimately support women's motivation and engagement as well as their compliance and adherence to the intervention, allowing them to use such digital solutions more freely while still giving the feeling of a more personalized experience.

Coherently, and in line with past evidence [60-62], women have expressed a desire for web-based solutions that support their psychological well-being. Furthermore, the data in this study seem to suggest that women might prefer app-based solutions, as almost all (3/4, 75%) participants that completed at least up to T6 used the app version instead of the web version to follow

XSL•FO
RenderX

the intervention. Compared with web-based programs, interventions deployed through a smartphone app are much more readily available wherever and whenever, thus being overall easier to use and access.

Limitations

This paper reports the first-phase study's findings, highlighting valuable insights and several limitations that can guide future research and the development of digital interventions. First and foremost is the inherent challenge of having endeavored to replicate or simulate an in-person intervention digitally, a task that resulted in an ineffective outcome in our case. It is crucial to acknowledge that digital solutions offer a unique environment and set of possibilities that distinguish them from face-to-face interactions. Attempting to mirror traditional methods within this digital landscape may fall short of fully capitalizing on the advantages that digital interventions can bring. With our findings, we thus strengthened the idea that digital interventions should not be mere replicas of their in-person counterparts; rather, they should harness distinctive strengths and capabilities. This challenge underscores the need for innovation and adaptation, as well as a recognition that a direct translation of traditional methods may not always yield optimal results in the digital sphere. Future studies should first try to conduct workshops with end users in which the in-person protocol is administered to discuss feasible changes to the intervention, thus singling out the intervention's main principle and appreciated practices and then adapting them to the digital format. Furthermore, a subsequent limitation of the study overall is the limited sample size, which has prevented the possibility of computing any statistical comparisons between the 2 groups as well as generalizing findings. Nonetheless, given the preliminary and exploratory nature of this study, aimed at setting the base for the co-design of the final intervention, the data collected were still able to provide valuable insights directing the refinement and future developments of the intervention. Further limitations are attributed to the platforms used, as they are created with different purposes than what they were used for in this study. This has warranted structuring the intervention content based on the functionalities of this platform instead of creating a tool tailored to the intervention requirements; therefore, future studies are advised to avoid using such tools to administer and evaluate digital intervention, even in the first-phase studies such as this one. However, it should still be emphasized that such platforms (ie, Moodle and Google Documents) have allowed to "prototype" the intervention in a time- and cost-efficient manner while collecting the information needed to refine the intervention and to create a specific app that can meet the requirements and preferences of the users.

Conclusions

This study had the purpose of evaluating the feasibility of the BATD-R protocol [31] structured as an internet-based self-help intervention among pregnant women with subclinical depression symptoms while comparing its guided versus unguided versions. A subsequent goal was the collection of women's feedback and perceptions of the intervention content through a prototyped version based on an e-learning platform to allow the co-design of the final intervention.

Overall, the ease with which the intervention can be followed has emerged as a central component to account for in the future developments of the intervention, in terms of the intervention itself as well as the platform used to administer it. The greater effort perceived was related to the homework, whereby women did emphasize that performing them was effortful and time-consuming. Even so, comparing the 2 versions of intervention, the guided version was more well-received than the unguided version. Nonetheless, both groups expressed satisfaction with the intervention's content and felt that it was relevant to their personal experiences with pregnancy.

Acknowledgments

The authors would like to thank the University of Padova and the Scuola di Specializzazione in Psicologia del Ciclo di Vita (Specialization School of Life Cycle Psychology, Padova, Italy) for allowing the inclusion of their psychotherapist in training to work as guides in the guided intervention. Moreover, the authors would like to thank the Artigianelli Institute (Trento, Italy) for creating the videos' illustrative content.

Conflicts of Interest

None declared.

References

1. American Psychiatric Association. Diagnostic And Statistical Manual Of Mental Disorders 5Ed Dsm-5. New York, NY: American Psychological Association; 2013.
2. Batt MM, Duffy KA, Novick AM, Metcalf CA, Epperson CN. Is postpartum depression different from depression occurring outside of the perinatal period? A review of the evidence. Focus (Am Psychiatr Publ) 2020 Apr;18(2):106-119 [FREE Full text] [doi: 10.1176/appi.focus.20190045] [Medline: 33162848]
3. Howard LM, Khalifeh H. Perinatal mental health: a review of progress and challenges. World Psychiatry 2020 Oct;19(3):313-327 [FREE Full text] [doi: 10.1002/wps.20769] [Medline: 32931106]
4. Yin X, Sun N, Jiang N, Xu X, Gan Y, Zhang J, et al. Prevalence and associated factors of antenatal depression: systematic reviews and meta-analyses. Clin Psychol Rev 2021 Mar;83:101932. [doi: 10.1016/j.cpr.2020.101932] [Medline: 33176244]

XSL•FO
RenderX

5. Wang Z, Liu J, Shuai H, Cai Z, Fu X, Liu Y, et al. Mapping global prevalence of depression among postpartum women. Transl Psychiatry 2021 Oct 20;11(1):543 [FREE Full text] [doi: 10.1038/s41398-021-01663-6] [Medline: 34671011]
6. Agostini F, Neri E, Salvatori P, Dellabartola S, Bozicevic L, Monti F. Antenatal depressive symptoms associated with specific life events and sources of social support among Italian women. Matern Child Health J 2015 May 11;19(5):1131-1141. [doi: 10.1007/s10995-014-1613-x] [Medline: 25303805]
7. Cena L, Mirabella F, Palumbo G, Gigantesco A, Trainini A, Stefana A. Prevalence of maternal antenatal and postnatal depression and their association with sociodemographic and socioeconomic factors: a multicentre study in Italy. J Affect Disord 2021 Jan 15;279:217-221 [FREE Full text] [doi: 10.1016/j.jad.2020.09.136] [Medline: 33069119]
8. Corbani IE, Rucci P, Iapichino E, Quartieri Bollani M, Cauli G, Ceruti MR, et al. Comparing the prevalence and the risk profile for antenatal depressive symptoms across cultures. Int J Soc Psychiatry 2017 Nov 14;63(7):622-631. [doi: 10.1177/0020764017725543] [Medline: 28805152]
9. Giardinelli L, Innocenti A, Benni L, Stefanini MC, Lino G, Lunardi C, et al. Depression and anxiety in perinatal period: prevalence and risk factors in an Italian sample. Arch Womens Ment Health 2012 Mar 29;15(1):21-30. [doi: 10.1007/s00737-011-0249-8] [Medline: 22205237]
10. Grussu P, Quatraro RM. Prevalence and risk factors for a high level of postnatal depression symptomatology in Italian women: a sample drawn from ante-natal classes. Eur Psychiatr 2020 Apr 16;24(5):327-333. [doi: 10.1016/j.eurpsy.2009.01.006]
11. Hadfield H, Wittkowski A. Women's experiences of seeking and receiving psychological and psychosocial interventions for postpartum depression: a systematic review and thematic synthesis of the qualitative literature. J Midwifery Womens Health 2017 Nov;62(6):723-736. [doi: 10.1111/jmwh.12669] [Medline: 29210501]
12. Baldisserotto ML, Miranda Theme M, Gomez LY, Dos Reis TB. Barriers to seeking and accepting treatment for perinatal depression: a qualitative study in Rio de Janeiro, Brazil. Community Ment Health J 2020 Jan 10;56(1):99-106. [doi: 10.1007/s10597-019-00450-4] [Medline: 31512080]
13. Fonseca A, Gorayeb R, Canavarro MC. Women s help-seeking behaviours for depressive symptoms during the perinatal period: socio-demographic and clinical correlates and perceived barriers to seeking professional help. Midwifery 2015 Dec;31(12):1177-1185. [doi: 10.1016/j.midw.2015.09.002] [Medline: 26433622]
14. Fonseca A, Ganho-Ávila A, Lambregtse-van den Berg M, Lupattelli A, de la Fé Rodriguez-Muñoz M, Ferreira P, et al. Emerging issues and questions on peripartum depression prevention, diagnosis and treatment: a consensus report from the cost action riseup-PPD. J Affect Disord 2020 Sep 01;274:167-173. [doi: 10.1016/j.jad.2020.05.112] [Medline: 32469800]
15. van den Heuvel JF, Groenhof TK, Veerbeek JH, van Solinge WW, Lely AT, Franx A, et al. eHealth as the next-generation perinatal care: an overview of the literature. J Med Internet Res 2018 Jun 05;20(6):e202 [FREE Full text] [doi: 10.2196/jmir.9262] [Medline: 29871855]
16. Hussain-Shamsy N, Shah A, Vigod SN, Zaheer J, Seto E. Mobile health for perinatal depression and anxiety: scoping review. J Med Internet Res 2020 Apr 13;22(4):e17011 [FREE Full text] [doi: 10.2196/17011] [Medline: 32281939]
17. Mancinelli E, Bassi G, Gabrielli S, Salcuni S. The efficacy of digital cognitive-behavioral interventions in supporting the psychological adjustment and sleep quality of pregnant women with sub-clinical symptoms: a systematic review and meta-analysis. Int J Environ Res Public Health 2022 Aug 03;19(15):9549 [FREE Full text] [doi: 10.3390/ijerph19159549] [Medline: 35954901]
18. Mancinelli E, Dell'Arciprete G, Pattarozzi D, Gabrielli S, Salcuni S. Digital behavioral activation interventions during the perinatal period: scoping review. JMIR Pediatr Parent 2023 Mar 28;6:e40937 [FREE Full text] [doi: 10.2196/40937] [Medline: 36853756]
19. Dimidjian S, Barrera M, Martell C, Muñoz RF, Lewinsohn PM. The origins and current status of behavioral activation treatments for depression. Annu Rev Clin Psychol 2011;7:1-38. [doi: 10.1146/annurev-clinpsy-032210-104535] [Medline: 21275642]
20. Dimidjian S, Goodman SH, Sherwood NE, Simon GE, Ludman E, Gallop R, et al. A pragmatic randomized clinical trial of behavioral activation for depressed pregnant women. J Consult Clin Psychol 2017 Jan;85(1):26-36 [FREE Full text] [doi: 10.1037/ccp0000151] [Medline: 28045285]
21. Jacobson NS, Martell CR, Dimidjian S. Behavioral activation treatment for depression: returning to contextual roots. Clin Psychol Sci Pract 2001;8(3):255-270. [doi: 10.1093/clipsy.8.3.255]
22. Soucy Chartier I, Provencher MD. Behavioural activation for depression: efficacy, effectiveness and dissemination. J Affect Disord 2013 Mar 05;145(3):292-299. [doi: 10.1016/j.jad.2012.07.023] [Medline: 22884236]
23. Kanter JW, Manos RC, Bowe WM, Baruch DE, Busch AM, Rusch LC. What is behavioral activation? A review of the empirical literature. Clin Psychol Rev 2010 Aug;30(6):608-620. [doi: 10.1016/j.cpr.2010.04.001] [Medline: 20677369]
24. ISO 9241-11:2018: ergonomics of human-system interaction: part 11: usability: definitions and concepts. International Organization for Standardization. 2018. URL: https://www.iso.org/standard/63500.html [accessed 2024-04-29]
25. O'Brien HL, Cairns P, Hall MA. A practical approach to measuring user engagement with the refined user engagement scale (UES) and new UES short form. Int J Hum Comput Stud 2018 Apr;112:28-39. [doi: 10.1016/j.ijhcs.2018.01.004]

XSL•FO
RenderX

26. Davis JA, Ohan JL, Gibson LY, Prescott SL, Finlay-Jones AL. Understanding engagement in digital mental health and well-being programs for women in the perinatal period: systematic review without meta-analysis. J Med Internet Res 2022 Aug 09;24(8):e36620 [FREE Full text] [doi: 10.2196/36620] [Medline: 35943773]
27. Vial S, Boudhraâ S, Dumont M. Human-centered design approaches in digital mental health interventions: exploratory mapping review. JMIR Ment Health 2022 Jun 07;9(6):e35591 [FREE Full text] [doi: 10.2196/35591] [Medline: 35671081]
28. Dorst K, Dijkhuis J. Comparing paradigms for describing design activity. Des Stud 1995 Apr;16(2):261-274. [doi: 10.1016/0142-694x(94)00012-3]
29. Dorst K. Design beyond design. She Ji J Des Econ Innov 2019;5(2):117-127. [doi: 10.1016/j.sheji.2019.05.001]
30. Dorst K. Design research: a revolution-waiting-to-happen. Des Stud 2008 Jan;29(1):4-11. [doi: 10.1016/j.destud.2007.12.001]
31. Lejuez CW, Hopko DR, Acierno R, Daughters SB, Pagoto SL. Ten year revision of the brief behavioral activation treatment for depression: revised treatment manual. Behav Modif 2011 Mar;35(2):111-161. [doi: 10.1177/0145445510390929] [Medline: 21324944]
32. Lyon AR, Koerner K. User-centered design for psychosocial intervention development and implementation. Clin Psychol (New York) 2016 Jun 17;23(2):180-200 [FREE Full text] [doi: 10.1111/cpsp.12154] [Medline: 29456295]
33. Czajkowski SM, Powell LH, Adler N, Naar-King S, Reynolds KD, Hunter CM, et al. From ideas to efficacy: the ORBIT model for developing behavioral treatments for chronic diseases. Health Psychol 2015 Oct;34(10):971-982 [FREE Full text] [doi: 10.1037/hea0000161] [Medline: 25642841]
34. Kroenke K, Spitzer RL, Williams JB. The PHQ-9: validity of a brief depression severity measure. J Gen Intern Med 2001 Sep;16(9):606-613 [FREE Full text] [doi: 10.1046/j.1525-1497.2001.016009606.x] [Medline: 11556941]
35. Randomizer. AnnaBet. URL: https://annabet.com/it/tools/Randomizer/ [accessed 2023-06-19]
36. Daughters SB, Braun AR, Sargeant MN, Reynolds EK, Hopko DR, Blanco C, et al. Effectiveness of a brief behavioral treatment for inner-city illicit drug users with elevated depressive symptoms: the life enhancement treatment for substance use (LETS Act!). J Clin Psychiatry 2008 Jan 15;69(1):122-129. [doi: 10.4088/jcp.v69n0116] [Medline: 18312046]
37. MacPherson L, Tull MT, Matusiewicz AK, Rodman S, Strong DR, Kahler CW, et al. Randomized controlled trial of behavioral activation smoking cessation treatment for smokers with elevated depressive symptoms. J Consult Clin Psychol 2010 Mar;78(1):55-61. [doi: 10.1037/a0017939] [Medline: 20099950]
38. Hopko DR, Lejuez CW, LePage JP, Hopko SD, McNeil DW. A brief behavioral activation treatment for depression. A randomized pilot trial within an inpatient psychiatric hospital. Behav Modif 2003 Sep 26;27(4):458-469. [doi: 10.1177/0145445503255489] [Medline: 12971122]
39. Sorveglianza dei disturbi depressivi: come lavora il Passi. Istituto Superiore di Sanità. URL: https://www.epicentro.iss.it/passi/DepressioneContesto [accessed 2023-01-29]
40. American Psychiatric Association. Diagnostic and Statistical Manual of Mental Disorders: Dsm-IV : International Version With Icd-10 Codes. New York, NY: American Psychiatric Association; 1998.
41. Shevlin M, Butter S, McBride O, Murphy J, Gibson-Miller J, Hartman TK, et al. Measurement invariance of the Patient Health Questionnaire (PHQ-9) and Generalized Anxiety Disorder scale (GAD-7) across four European countries during the COVID-19 pandemic. BMC Psychiatry 2022 Mar 01;22(1):154 [FREE Full text] [doi: 10.1186/s12888-022-03787-5] [Medline: 35232409]
42. Cox JL, Holden JM, Sagovsky R. Detection of postnatal depression. Development of the 10-item Edinburgh postnatal depression scale. Br J Psychiatry 1987 Jun 2;150(06):782-786. [doi: 10.1192/bjp.150.6.782] [Medline: 3651732]
43. Benvenuti P, Ferrara M, Niccolai C, Valoriani V, Cox JL. The Edinburgh Postnatal Depression Scale: validation for an Italian sample. J Affect Disord 1999 May;53(2):137-141. [doi: 10.1016/s0165-0327(98)00102-5] [Medline: 10360408]
44. Spitzer RL, Kroenke K, Williams JB, Löwe B. A brief measure for assessing generalized anxiety disorder: the GAD-7. Arch Intern Med 2006 May 22;166(10):1092-1097. [doi: 10.1001/archinte.166.10.1092] [Medline: 16717171]
45. Cohen S, Kamarck T, Mermelstein R. A global measure of perceived stress. J Health Soc Behav 1983 Dec;24(4):385. [doi: 10.2307/2136404]
46. Mondo M, Sechi C, Cabras C. Psychometric evaluation of three versions of the Italian Perceived Stress Scale. Curr Psychol 2019 Jan 08;40(4):1884-1892. [doi: 10.1007/s12144-019-0132-8]
47. Manos RC, Kanter JW, Luo W. The behavioral activation for depression scale-short form: development and validation. Behav Ther 2011 Dec;42(4):726-739. [doi: 10.1016/j.beth.2011.04.004] [Medline: 22036000]
48. Armento ME, Hopko DR. The Environmental Reward Observation Scale (EROS): development, validity, and reliability. Behav Ther 2007 Jun;38(2):107-119. [doi: 10.1016/j.beth.2006.05.003] [Medline: 17499078]
49. Domnich A, Arata L, Amicizia D, Signori A, Patrick B, Stoyanov S, et al. Development and validation of the Italian version of the mobile application rating scale and its generalisability to apps targeting primary prevention. BMC Med Inform Decis Mak 2016 Jul 07;16(1):83 [FREE Full text] [doi: 10.1186/s12911-016-0323-2] [Medline: 27387434]
50. RStudio Team. RStudio: integrated development for R. RStudio. URL: https://www.r-project.org/conferences/useR-2011/abstracts/180111-allairejj.pdf [accessed 2024-04-29]
51. Braun V, Clarke V. Using thematic analysis in psychology. Qual Res Psychol 2006 Jan;3(2):77-101. [doi: 10.1191/1478088706qp063oa]

XSL•FO
RenderX

52. Gabrielli S, Rizzi S, Carbone S, Donisi V. A Chatbot-based coaching intervention for adolescents to promote life skills: pilot study. JMIR Hum Factors 2020 Mar 14;7(1):e16762 [FREE Full text] [doi: 10.2196/16762] [Medline: 32130128]
53. World Medical Association. Dichiarazione di Helsinki della World Medical Association: principi etici per la ricerca biomedica che coinvolge gli esseri umani. Evidence 2013;5(10):e1000059. [doi: 10.4470/E1000059]
54. Hayes SC, Hofmann SG. "Third-wave" cognitive and behavioral therapies and the emergence of a process-based approach to intervention in psychiatry. World Psychiatry 2021 Oct 09;20(3):363-375 [FREE Full text] [doi: 10.1002/wps.20884] [Medline: 34505370]
55. Feliu-Soler A, Cebolla A, McCracken LM, D'Amico F, Knapp M, López-Montoyo A, et al. Economic impact of third-wave cognitive behavioral therapies: a systematic review and quality assessment of economic evaluations in randomized controlled trials. Behav Ther 2018 Jan;49(1):124-147 [FREE Full text] [doi: 10.1016/j.beth.2017.07.001] [Medline: 29405918]
56. Hayes SC, Hofmann SG. The third wave of cognitive behavioral therapy and the rise of process-based care. World Psychiatry 2017 Oct;16(3):245-246 [FREE Full text] [doi: 10.1002/wps.20442] [Medline: 28941087]
57. Nielsen J. Usability Engineering. Boston, MA: Academic Press; 1993.
58. Provoost S, Lau HM, Ruwaard J, Riper H. Embodied conversational agents in clinical psychology: a scoping review. J Med Internet Res 2017 May 09;19(5):e151 [FREE Full text] [doi: 10.2196/jmir.6553] [Medline: 28487267]
59. Gaffney H, Mansell W, Tai S. Conversational agents in the treatment of mental health problems: mixed-method systematic review. JMIR Ment Health 2019 Oct 18;6(10):e14166 [FREE Full text] [doi: 10.2196/14166] [Medline: 31628789]
60. Ashford MT, Ayers S, Olander EK. Interest in web-based treatments for postpartum anxiety: an exploratory survey. J Reprod Infant Psychol 2017 Sep 02;35(4):394-409. [doi: 10.1080/02646838.2017.1320364] [Medline: 29517373]
61. Lackie ME, Parrilla JS, Lavery BM, Kennedy AL, Ryan D, Shulman B, et al. Digital health needs of women with postpartum depression: focus group study. J Med Internet Res 2021 Jan 06;23(1):e18934 [FREE Full text] [doi: 10.2196/18934] [Medline: 33404506]
62. Wenze S, Miers QA, Battle CL. Postpartum mental health care for mothers of multiples: a qualitative study of new mothers' treatment preferences. J Psychiatr Pract 2020 May;26(3):201-214. [doi: 10.1097/PRA.0000000000000469] [Medline: 32421291]

Abbreviations

BA: behavioral activation
BATD-R: Behavioral Activation Treatment for Depression-Revised
DSM-IV: Diagnostic and Statistical Manual of Mental Disorders (Fourth Edition)
MARS: Mobile Application Rating Scale
PHQ-9: Patient Health Questionnaire-9
UE: user engagement
UX: user experience

Edited by Y Senathirajah; submitted 19.06.23; peer-reviewed by E Neter, E Toki; comments to author 02.09.23; revised version received 05.10.23; accepted 01.03.24; published 16.05.24.

Please cite as:
Mancinelli E, Gabrielli S, Salcuni S
A Digital Behavioral Activation Intervention (JuNEX) for Pregnant Women With Subclinical Depression Symptoms: Explorative Co-Design Study
JMIR Hum Factors 2024;11:e50098
URL: https://humanfactors.jmir.org/2024/1/e50098
doi:10.2196/50098
PMID:38753421

©Elisa Mancinelli, Silvia Gabrielli, Silvia Salcuni. Originally published in JMIR Human Factors (https://humanfactors.jmir.org), 16.05.2024. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on https://humanfactors.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Trust in and Acceptance of Artificial Intelligence Applications in Medicine: Mixed Methods Study

Daria Shevtsova[1,2], MSc; Anam Ahmed[1], MSc; Iris W A Boot[1], MSc; Carmen Sanges[3], PhD; Michael Hudecek[3], PhD; John J L Jacobs[4], PhD; Simon Hort[5], MSc; Hubertus J M Vrijhoef[1], PhD

[1]Panaxea bv, Den Bosch, Netherlands
[2]Vrije Universiteit Amsterdam, Amsterdam, Netherlands
[3]Universitätsklinikum Würzburg, Würzburg, Germany
[4]Ortec bv, Zoetermeer, Netherlands
[5]Fraunhofer Institute for Production Technology, Aachen, Germany

Corresponding Author:
Anam Ahmed, MSc
Panaxea bv
Pettelaarpark 84
Den Bosch, 5216 PP
Netherlands
Phone: 31 639421854
Email: anam.ahmed@panaxea.eu

Abstract

Background: Artificial intelligence (AI)–powered technologies are being increasingly used in almost all fields, including medicine. However, to successfully implement medical AI applications, ensuring trust and acceptance toward such technologies is crucial for their successful spread and timely adoption worldwide. Although AI applications in medicine provide advantages to the current health care system, there are also various associated challenges regarding, for instance, data privacy, accountability, and equity and fairness, which could hinder medical AI application implementation.

Objective: The aim of this study was to identify factors related to trust in and acceptance of novel AI-powered medical technologies and to assess the relevance of those factors among relevant stakeholders.

Methods: This study used a mixed methods design. First, a rapid review of the existing literature was conducted, aiming to identify various factors related to trust in and acceptance of novel AI applications in medicine. Next, an electronic survey including the rapid review–derived factors was disseminated among key stakeholder groups. Participants (N=22) were asked to assess on a 5-point Likert scale (1=irrelevant to 5=relevant) to what extent they thought the various factors (N=19) were relevant to trust in and acceptance of novel AI applications in medicine.

Results: The rapid review (N=32 papers) yielded 110 factors related to trust and 77 factors related to acceptance toward AI technology in medicine. Closely related factors were assigned to 1 of the 19 overarching umbrella factors, which were further grouped into 4 categories: human-related (ie, the type of institution AI professionals originate from), technology-related (ie, the explainability and transparency of AI application processes and outcomes), ethical and legal (ie, data use transparency), and additional factors (ie, AI applications being environment friendly). The categorized 19 umbrella factors were presented as survey statements, which were evaluated by relevant stakeholders. Survey participants (N=22) represented researchers (n=18, 82%), technology providers (n=5, 23%), hospital staff (n=3, 14%), and policy makers (n=3, 14%). Of the 19 factors, 16 (84%) human-related, technology-related, ethical and legal, and additional factors were considered to be of high relevance to trust in and acceptance of novel AI applications in medicine. The patient's gender, age, and education level were found to be of low relevance (3/19, 16%).

Conclusions: The results of this study could help the implementers of medical AI applications to understand what drives trust and acceptance toward AI-powered technologies among key stakeholders in medicine. Consequently, this would allow the implementers to identify strategies that facilitate trust in and acceptance of medical AI applications among key stakeholders and potential users.

(JMIR Hum Factors 2024;11:e47031) doi:10.2196/47031

XSL•FO
RenderX

KEYWORDS

trust; acceptance; artificial intelligence; medicine; mixed methods; rapid review; survey

Introduction

Artificial Intelligence

Artificial intelligence (AI) is commonly defined as a computer system that uses statistical models, diverse algorithms, and self-modifying systems to make predictions and decisions based on its own aggregated experience. It can therefore perform tasks that usually require or even surpass the human level of intelligence [1,2]. AI has been increasingly integrated in the health care sector, where it helps with administrative workflows, diagnostic image analysis, robotic surgery, and clinical decision-making. Consequently, medical AI applications allow, amongst other things, earlier disease detection, patient-tailored treatments, and more efficient follow-ups, which should drive the health care costs down upon implementation [3]. Although medical AI applications provide various advantages to the current health care system, such as increased efficiency and improved workflows [4], there are also various challenges associated with AI implementation. For instance, a large share of potential users has concerns over privacy issues [5,6]. Equity and fairness are other important concerns, since there is a risk of perpetuating bias within data sets being adopted by AI technology [5,7,8]. Further, the implementation of AI into medical practice raises the question of accountability, since it is currently unclear whether technology developers, hospitals, or regulators should be responsible for mistakes or undesirable outcomes from the use of an AI application.

Trust and Acceptance

To ensure successful implementation of medical AI applications, it is essential to build trust in and acceptance of AI technology among its users [9,10]. In this study, a model of key drivers of trust in and acceptance of AI systems was used [11]. According to this model, trust is influenced by 4 drivers: current safeguards, job impact of AI, familiarity of AI, and AI uncertainty. Current safeguards indicate the belief that current regulations and laws are adequate for ensuring the safety of AI and the protection of people who use it. Job impact refers to the belief that there will be more jobs generated than eliminated due to AI implementation. Familiarity with AI is the level of understanding of how AI technology works and how AI applications are used. These 3 drivers have a positive influence on trust, with current safeguards being its strongest driver. The fourth driver, AI uncertainty, impacts trust in a negative way. It implies the belief that the impact of AI on society is unpredictable and the technology is still not fully explored. Overall, these drivers influence the extent to which people trust the AI system and believe it to be trustworthy. Trust, then, is a large contributor to the level of acceptance, which is the extent to which people accept or approve of AI and are willing to use it without resistance [11]. In the scientific literature, trust can be defined in different ways. In this study, we used literature-derived definitions of trust and acceptance in the context of AI implementation, namely:

- Trust is the belief of an individual that an AI application will do what it promises [12,13].
- Acceptance is the willingness of an individual to use the AI application in medicine [14].

Therefore, it can be argued that acceptance of an AI application depends on trust people have toward this technology [11,15,16]. At the same time, people can often accept their usage of technologies without necessarily trusting them [17]. Therefore, it is important to consider the 2 concepts separately as well as together.

Overall, widespread trust in and acceptance of an AI application is crucial for successful introduction and implementation of the technology. Failure to ensure trust in and acceptance of AI technology would pose the risk of “stifling innovation” and causing unnecessary “opportunity costs” [18]. The lack of trust in AI applications in medicine impedes their adoption in health care, compounded by inadequate public assurance and attention to concerns, thereby exacerbating these challenges. In addition, the anticipated benefits of AI-based innovations can coexist with significant acceptance barriers [15,18-21].

Investigating what factors contribute to trust in and acceptance of AI technology in medicine would help us understand how to make the implementation and regulatory approval of AI-powered advanced therapy manufacturing systems as efficient as possible. This can be achieved by collecting insights into stakeholders’ perspectives with regard to trust and acceptance toward medical AI applications [2,22]. Factors contributing to trust and acceptance toward medical AI applications can be attributed a different weight by various groups of stakeholders with distinct roles in AI.

Study Objectives

Since AI applications are still relatively new, users and providers are hesitant to trust and accept this new technology without restrictions. As for the future implementation of AI applications in treatment centers, it is essential that stakeholders (eg, clinicians, researchers, hospital staff) accept and trust the innovative AI-based manufacturing platform. Therefore, the aim of this study was first to identify the factors related to trust in and acceptance of AI technology in medicine and second to assess the relevance of those factors among relevant stakeholders in medicine.

Methods

Study Setting

This study is part of the European Union’s (EU) Horizon 2020 project AIDPATH (AI-driven Decentralized Production for Advanced Therapies in the Hospital; grant agreement number 101016909) [22,23]. It is an upcoming state-of-the-art AI application in hospitals, which aims to develop an AI-driven, automated chimeric antigen receptor T cell (CAR-T) manufacturing platform at the point of care as a treatment for acute leukemia and lymphoma. In CAR-T therapy, the patient’s

XSL•FO
RenderX

own T cells are removed, genetically modified, and reinfused into the patient in order to find and eliminate tumour cells. Current production is characterized by laborious manual process steps, complex logistics, and a lack of process understanding. This results in long delivery times (up to 21 days) and high costs (approx €320,000, or US $347,890, per treatment) [24,25]. For this reason, AIDPATH is developing a system to fully automate the manufacturing process, from the provision of patient cells to the injection directly in the hospital. An important building block for effective and equitable manufacturing is AI. AI can provide essential process insights into the cell's characteristics and behavior. This offers a significant benefit for adaptive control of the whole process and the design of personalized process protocols. Furthermore, AI can assist cost-effective platform operation in a smart manufacturing hospital by improving manufacturing schedules and resource management [26]. In general, successful implementation of AIDPATH would serve as an example of an effective AI technology that automates the production and delivery of advanced therapy medicinal products (ATMPs). Furthermore, AI-powered technology can form the basis for a deployable platform for further pilot trials in multiple hospitals and would create a model innovation system for smart manufacturing hospitals [2,22].

In this study, to meet the study objectives, a rapid literature review was conducted, followed by a survey.

Rapid Literature Review

A rapid literature review of peer- and non-peer-reviewed publications was conducted to identify factors related to trust in and acceptance of AI applications used in medicine. As an alternative method to systematic reviews, a rapid review allows for accelerated synthesis of up-to-date evidence, while efficiently informing latest findings in recent health care research [27]. The peer- and non-peer-reviewed literature needed to be published between 2012 and 2022 in English. Data on attitudes toward AI in relation to prognosis, diagnosis, treatment, and care were included. The search was performed in PubMed/MEDLINE with the following search syntax: ((trust) OR (acceptance) OR (attitude) OR (perspective) OR (perception)) AND ((AI) OR (artificial intelligence) OR (machine learning) OR (deep learning)) AND (((prognosis) OR (diagnosis) OR (treatment) OR (care)) OR ((medic*) OR (clinic*) OR (hospital) OR (smart hospital) OR (health care)) AND ((survey) OR (questionnaire) OR (interview)). The reason for inclusion of only survey-, questionnaire-, or interview-based research in the search terms was due to their direct relevance to our research objectives.

In the non-peer-reviewed literature search, similar terms and time frame of publication were used and the first 10 pages on the Google Search engine were examined to identify other relevant papers and reports by (non)governmental and research organizations. This allowed the study findings to be applicable to a broad range of medical AI applications. Papers were screened, and data were extracted by 2 authors (DS and AA). The selected literature was analyzed to identify key trends and explanatory factors related to trust and acceptance toward medical AI applications. The factors were then grouped into 4 categories: human-related, technology-related, legal and ethical, and additional factors. These factor groups formed the basis of the survey designed to investigate factor relevance. This was performed independently by 2 authors (DS and AA).

Survey

The survey was reported in accordance with the CHERRIES (Checklist for Reporting Results of Internet E-Surveys) guidelines [28]. The survey in English assessed the relevance of the factors related to trust in and acceptance of novel AI applications in medicine. The survey started with an introduction to AI applications in medicine and AIDPATH, followed by 7 general questions on each participant's background, including gender, age, the country they worked in, years of experience, the stakeholder group they belonged to, their familiarity with AI applications in medicine, and their general view on AI. In the last question, the following distinction was made between the answer options: "I embrace AI" meant welcoming and using AI as a constituent part of their work or life, "I approve of AI" implied that the participant agreed with the use of AI in their work or life but did not use it themselves, and "I accept AI" referred to acknowledging the use of AI in work or life but not being ready to fully approve it.

In the core section of the survey, the definitions of trust and acceptance were provided as a reference for participants. The core part also consisted of 2 identical lists of 19 factors related to trust and acceptance toward AI applications in medicine. Each factor was categorized into human-related, technology-related, legal and ethical, or additional factors. Human-related factors were linked to AI professionals assessed the relevance of the type of organization the AI professionals were affiliated to and the purpose to innovate with a specific AI application. With respect to health care professionals, the factors were related to the knowledge of AI applications and the attitude toward AI application usage in medicine. In relation to patients, the relevance of the following factors was assessed: general knowledge of AI applications in medicine, the attitude toward AI application usage in medicine, and the patient's age, gender, and level of education. Furthermore, participants were asked to evaluate the relevance of transparency between all parties involved in AI application use. Technology-related factors related to the performance of AI applications in medicine, the possibility of their integration into existing clinical workflows, a clear balance of risks and benefits of the AI applications, and the explainability and transparency of processes and outcomes. The legal and ethical factors were related to the adequacy of regulations and governance of AI applications in medicine, data use transparency, and clear accountability and responsibility for an AI application. The additional factors were concerned with the environmental sustainability of AI applications and AI's impact on job availability. For each factor, participants could indicate each factor's relevance to trust in and acceptance of AI applications from their stakeholder perspective using a Likert scale of 1-5, where 1 stood for "not relevant," 3 for "not irrelevant, nor relevant," and 5 for "relevant." Throughout the survey, "relevant" meant being highly significant for ensuring trust in or acceptance of AI applications, while "irrelevant" meant no significance. The N/A (not applicable) option was available as well for each factor. Open questions at the end of both sections

XSL•FO
RenderX

allowed participants to suggest other relevant factors related to trust in or acceptance of AI applications that were not mentioned in the survey. Furthermore, the participants were invited to suggest any other factors, different from trust, deemed important for acceptance of AI applications in medicine.

Sampling

Using the convenience sampling method [29], AIDPATH Consortium members were requested to invite stakeholders in their network but outside the AIDPATH Consortium to fill in the survey on the SurveyMonkey platform. The survey was distributed by email to members of relevant stakeholder groups to capture their professional perspectives (eg, clinicians, scientists, and policy makers). Data were collected from April to May 2022 and analyzed using Microsoft Excel.

Data Collection and Analysis

After participants were asked to rate the relevance of each factor from 1 (irrelevant) to 5 (relevant), the mean score of each factor was determined by assigning each response a weight from 1 to 5. Next, means scores were calculated by finding an average of the sum of response values for each question. To visualize the survey responses and compare the mean scores for each factor included in the survey, a spider diagram was charted. This provided an overview of the factors' relevance and their relative importance in influencing both trust and acceptance toward AI applications in medicine. In addition, a scatter plot was created to obtain an overview of the interrelationship between the relevance to trust (x axis) in and acceptance (y axis) of AI applications in medicine. The plot allowed us to identify the degree of relevance of each factor in relation to both trust and acceptance. To classify the factors based on their relevance, score ranges were established. Factors with mean scores from 1 to 3 were considered to be of low relevance, while factors from 4 to 5 were deemed of high relevance. The open-question responses were considered when interpreting numerical data.

Ethical Considerations

Under Dutch law, no ethical approval was required according to Article 1b of the Dutch Medical Research in Human Subjects Act [30]. However, all participants were informed about the study objectives, their verbal consent was obtained, and all data were processed anonymously. All responses were recorded anonymously. Participants were informed of their right to withdraw from the study at any time without any consequences. They were not financially compensated.

Results

Rapid Literature Review

The literature search (Figure 1) yielded 301 hits in the PubMed database and 105 hits through gray literature search and snowballing. After screening titles and abstracts, 284 (70%) records were excluded. After full-text screening, 90 (73.8%) records were excluded primarily due to the absence of concepts of trust or acceptance and a lack of factors related to trust or acceptance in the main text or data-containing figures. As a result, 32 (26.2%) papers and reports [7,9-12,15,16,19,21,31-53] were included in the data analysis.

Overall, the rapid review identified a total of 110 factors related to trust and 77 factors related to acceptance toward medical AI technology. The full list of factors identified through the rapid review with corresponding studies can be found in Multimedia Appendix 1. Tables 1-4 show all factors from the rapid review, each with the frequency of its appearance in the literature and the corresponding overarching umbrella factors. Some factors from a single study are repeated in the same category in Tables 1 and 2 on trust and Tables 3 and 4 on acceptance or within the same category (eg, health care professionals and patients subsections of the human-related factors section). The most frequently reported human-related factors related to trust (Tables 1 and 2) in medical AI applications were knowledge and understanding of AI by health care professionals and knowledge and education of AI among patients. In terms of technology-related factors, accuracy, transparency, reliability, safety, and explainability of medical AI applications and their functioning appeared most often in the literature. Regarding legal and ethical factors, the most frequently occurring factors included fairness and equity of medical AI technology and the privacy and security of personal data handled by the AI systems. The most frequently presented human-related factors related to acceptance (Tables 3 and 4) of medical AI technology were the perceived usefulness and provision of better medical services by the AI technology. Regarding technology-related factors linked to acceptance, performance expectancy, design and output quality, and transparency were stated in the literature most often. A wide range of legal and ethical factors were mentioned in the literature, including adequate regulations of medical AI technology, protection and security of patients' data, and the allocation of accountability and responsibility for the (mal)functioning of an AI application. There were additional factors related to trust in and acceptance of medical AI technology (Tables 1-4). These included replacement of doctors by machines that lack a human touch and moral support, labor market implications, and environmental sustainability. Three studies also highlighted that acceptance of a medical AI application is directly related to trust in the AI application. Overall, there were fewer factors related to acceptance than those related to trust, whereas most of the overarching umbrella factors were fully represented in both tables. Therefore, an identical list of umbrella factors allocated within the 4 categories (human-related, technology-related, legal and ethical, and additional factors) was used in the survey for investigating the relevance of factors for both trust in and acceptance of AI applications in medicine.

XSL•FO
RenderX

Figure 1. PRISMA flow diagram of the rapid review literature screening. AI: artificial intelligence.

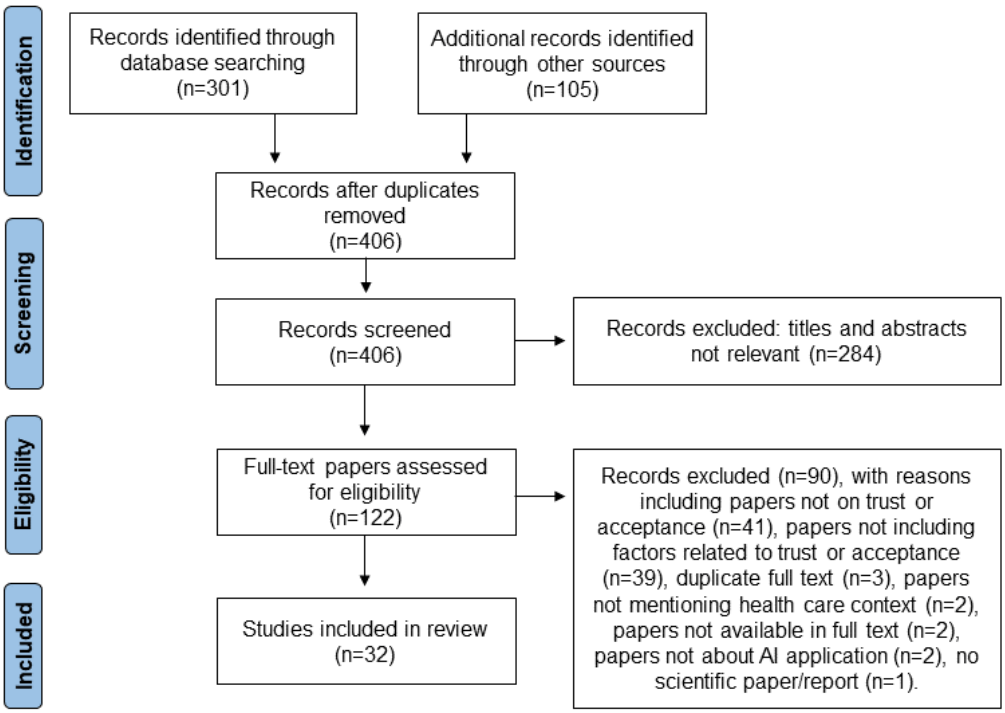


Table 1. Human-related factors related to trust (N=110) in medical AI[a] applications (22/32, 68.8%, studies).

Factor category and factors from the rapid review	Umbrella factors used in the survey
AI professionals	
AI company/provider (n=2, 9.1%)	Type of institution/organization of AI professionals (eg, university, technology company, commercial organization)
AI role (n=1, 4.5%); perceived helpfulness (n=1, 4.5%)	The purpose to innovate with a specific AI application in medicine (eg, financial vs societal)
Health care professionals	
Knowledge and understanding of AI (n=6, 27.3%); education (n=3, 13.6%)	Knowledge of AI applications in medicine (eg, by means of training and education)
Expectation of AI (n=1, 4.5%); perceived actionability (ie, clear recommendation for action; n=1, 4.5%); user's social network (n=1, 4.5%); user's media consumption (n=1, 4.5%)	Attitude toward AI application usage in medicine (eg, agreeableness, openness, conscientiousness, engagement)
Patients informed about AI application usage in the hospital	
Knowledge/education about AI (n=5, 22.7%); awareness of AI (n=2, 9.1%)	General knowledge of AI applications in medicine
Openness (to AI health care technologies and to judgments of potential benefits and harms; n=1, 4.5%); perceived benefit and lower concern (n=1, 4.5%); user's social network (n=1, 4.5%); user's media consumption (n=1, 4.5%)	Attitude toward AI application usage in medicine (eg, agreeableness, openness, conscientiousness)
Gender (n=2, 9.1%); age (n=1, 4.5%); type of user (n=1, 4.5%)	Age, gender, level of education
All parties	
Clinicians and patients interaction during AI integration (n=1, 4.5%); human agency and oversight (n=1, 4.5%)	Transparency between all involved parties (AI professionals, health care professionals, patients)

[a]AI: artificial intelligence.

XSL•FO
RenderX

Table 2. Other factors related to trust (N=110) in medical AI[a] applications (22/32, 68.8%, studies).

Factor category and factors from the rapid review	Umbrella factors used in the survey
Technology-related factors	
Accuracy (n=7, 31.8%); reliability (n=5, 22.7%); safety (n=4, 18.2%); design and output quality (n=2, 9.1%); performance expectancy (n=2, 9.1%); ability (n=1, 4.5%); perceived functionality (n=1, 4.5%); self-efficacy (n=1, 4.5%); tool itself (n=1, 4.5%)	Performance of AI applications in medicine (reproducibility of outcomes, accuracy)
Auditability (n=1, 4.5%); customizability (n=1, 4.5%); understandability (n=1, 4.5%); ease of integration into clinical workflows (n=1, 4.5%); convenience of use (n=1, 4.5%); usability (n=1, 4.5%); (over)alerting and excessive false-positive rate (n=1, 4.5%)	Possibility of integration of AI applications into existing clinical workflows
Risk and impact mitigation (n=1, 4.5%)	Clear balance of risks and benefits of the AI application
Transparency (n=6, 27.3%); explainability (n=5, 22.7%); evidence strength (n=2, 9.1%); benevolence (n=2, 9.1%); complexity (n=2, 9.1%); interpretability (n=2, 9.1%); integrity (n=1, 4.5%); predictability (n=1, 4.5%); trialability (n=1, 4.5%); trustworthiness (n=1, 4.5%)	Explainability and transparency of the processes and outcomes
Legal and ethical factors	
Fairness and equity (n=8, 36.4%); adequate regulations, legislation, and governance (n=3, 13.6%); ethical/legal implications (n=1, 4.5%)	Adequacy of the regulations and governance of AI applications in medicine
Personal data privacy and security (n=8, 36.4%); data used to train AI/cognitive bias (n=2, 9.1%); data sensitivity (n=1, 4.5%); respect and preservation of human dignity (n=1, 4.5%)	Data use transparency
Accountability (n=3, 13.6%); power-control balance (n=1, 4.5%)	Clear accountability and responsibility of the AI application (machine vs human responsibility)
Additional factors	
Environmental sustainability (n=1, 4.5%)	Environment-friendly AI application
Replacement of doctor/lack of human touch and moral support when evaluated by AI alone (n=1, 4.5%); labor market implications (n=1, 4.5%)	Impact on job availability (machines replacing humans)

[a]AI: artificial intelligence.

XSL•FO
RenderX

Table 3. Human-related factors related to acceptance (N=77) of medical AI[a] applications (14/32, 43.8%, studies).

Factor category and factors from the rapid review	Umbrella factors used in the survey
AI professionals	
AI company/provider (n=1, 7.1%); brand impact (n=1, 7.1%)	Type of institution/organization of AI professionals (eg, university, technology company, commercial organization)
Perceived usefulness (n=3, 21.4%); better medical services/ understanding of disease (n=3, 21.4%); improve the quality of people's lives (n=2, 14.3%); medical costs (n=2, 14.3%); AI role (eg, saving patients' time; n=1, 7.1%); miniaturization of hardware (n=1, 7.1%)	Purpose to innovate with a specific AI application in medicine (eg, financial vs societal)
Health care professionals	
Knowledge and understanding of AI (n=1, 7.1%)	Knowledge of AI applications in medicine (eg, by means of training and education)
Behavioral intention to use (n=2, 14.3%); effort expectancy (n=2, 14.3%); perceived ease of use (n=2, 14.3%); perceived usefulness (n=2, 14.3%); intrinsic motivation (n=1, 7.1%); interest in AI (n=1, 7.1%); professional identity (n=1, 7.1%); concerns about benefit to patient care (n=1, 7.1%); general impression of AI (n=1, 7.1%)	Attitude toward AI application usage in medicine (eg, agreeableness, openness, conscientiousness, engagement)
Patients informed about AI application usage in the hospital	
Knowledge/education about AI (n=1, 7.1%); awareness of AI (n=1, 7.1%)	General knowledge of AI applications in medicine
Behavioral intention to use (n=2, 14.3%); general impression (n=1, 7.1%); Interest in topic (n=1, 7.1%)	Attitude toward AI application usage in medicine (eg, agreeableness, openness, conscientiousness)
Age (n=1, 7.1%)	Age
All parties	
Expectations of others (n=2, 14.3%)	Transparency between all involved parties (AI professionals, healthcarehealth care professionals, patients)

[a]AI: artificial intelligence.

Table 4. Other factors related to acceptance (N=77) of medical AI[a] applications (14/32, 43.8%, studies).

Factor category and factors from the rapid review	Umbrella factors used in the survey
Technology-related factors	
Performance expectancy (n=4, 28.6%); design and output quality (n=4, 28.6%); accuracy (n=2, 14.3%); efficiency (n=1, 7.1%)	Performance of AI applications in medicine (reproducibility of outcomes, accuracy)
Perceived ease of use (n=2, 14.3%); user-friendliness (n=2, 14.3%); actual system use (n=1, 7.1%); compatibility (n=1, 7.1%); facilitating conditions (n=1, 7.1%)	Possibility of integration of AI applications into existing clinical workflows
Perceived risk (n=1, 7.1%)	Clear balance of risks and benefits of the AI application
Transparency (n=3, 21.4%); explainability (n=2, 14.3%); evidence strength (n=1, 7.1%); trustworthiness (n=1, 7.1%)	Explainability and transparency of the processes and outcomes
Legal and ethical factors	
Adequate regulations, legislation and governance (n=2, 14.3%); ethical risks (n=1, 7.1%); political support (n=1, 7.1%)	Adequacy of the regulations and governance of AI applications in medicine
Data protection/security (n=2, 14.3%); patients' consent to the continuous collection and processing of data (n=1, 7.1%)	Data use transparency
Accountability and responsibility (n=2, 14.3%); tort liability (n=1, 7.1%)	Clear accountability and responsibility of the AI application (machine vs human responsibility)
Additional factors	
Replacement of doctor/lack of human touch and moral support when evaluated by AI alone (n=1, 7.1%)	Impact on job availability (machines replacing humans)
Trust in AI applications (n=3, 21.4%)	Acceptance emerging from trust

[a]AI: artificial intelligence.

XSL•FO
RenderX

Survey

Participants

A total of 22 respondents participated in the survey, of which 18 (82%) completed the questions on trust and 15 (68%) completed the questions on acceptance. No reasons were provided for not completing the survey. Table 5 shows the characteristics of the survey participants, the majority (n=21, 95%) of whom came from European countries, were aged from 40 to 60 years, and had 0-10 or 21-30 years of professional experience.

Participants were mainly slightly (n=7, 32%) or moderately (n=8, 36%) familiar with AI-based devices used for clinical purposes (Figure 2). In thinking about AI, 9 (41%) of the participants indicated that the statement “I accept AI” best represents their view, followed by “I approve of AI” (n=6, 27%) and “I embrace AI” (n=5, 23%); see Figure 3.

Table 5. Characteristics of the participants (N=22).

Characteristic and type of participant	Participants, n (%)
Stakeholder group[a]	
Researchers	18 (82)
Technology providers	5 (23)
Hospital staff	3 (14)
Policy makers	3 (14)
Gender	
Female	8 (36)
Male	13 (59)
Prefer not to say	1 (5)
Age (years)	
≤30	2 (9)
31-39	3 (14)
40-49	6 (27)
50-59	5 (23)
≥60	6 (27)
Country of work	
Netherlands	11 (50)
Germany	3 (14)
Ireland	2 (9)
Spain	2 (9)
France	1 (5)
Hungary	1 (5)
India	1 (5)
Italy	1 (5)
Years of professional experience	
0-10	6 (27)
11-20	4 (18)
21-30	7 (32)
31-40	5 (23)

[a]Participants sometimes represented more than 1 stakeholder group.

XSL•FO
RenderX

Figure 2. Familiarity with AI-based devices used for clinical purposes (N=22). AI: artificial intelligence.

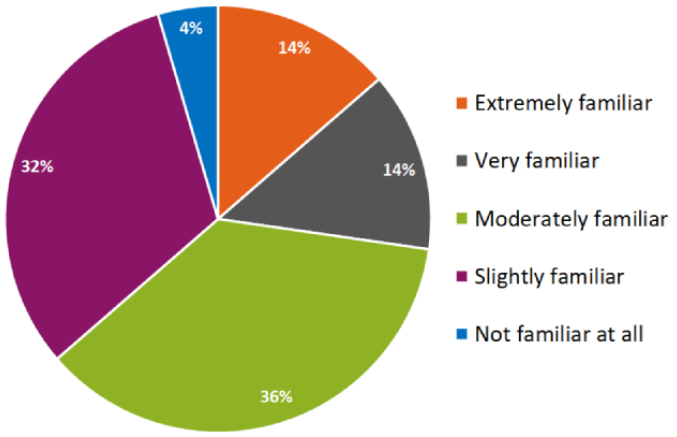
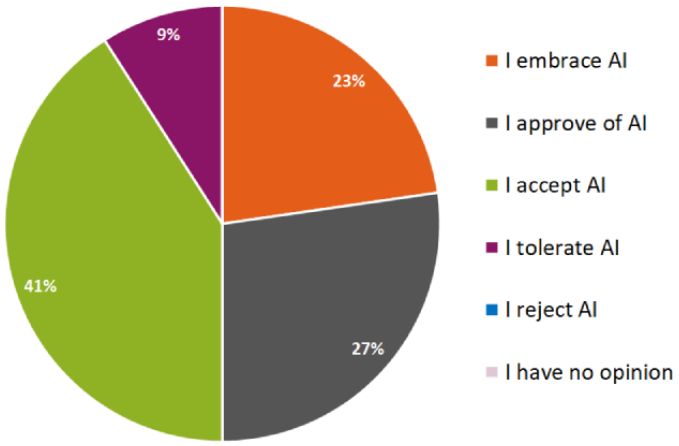


Figure 3. Statement best representing participants' view when thinking about AI (N=22). AI: artificial intelligence.



Relevance of Factors for Trust in and Acceptance of AI

In Table 6, the mean scores per factor for its relevance to trust and acceptance are shown. Figure 4 demonstrates a spider diagram with the 19 summarized statements and the corresponding mean scores of relevance to trust in and acceptance of AI applications in medicine. The degrees of relevance of the factors related to trust and to acceptance closely followed each other for all but 1 (5.3%) of the 19 factors. Only the type of AI organization was slightly more relevant to trust than to acceptance toward AI applications in medicine. In Figure 5, a scatter plot displays the combined relevance of the factors related to trust (x axis) and acceptance (y axis) toward medical AI applications. Of the 19 factors included in the survey, 3 (16%) were found to have, on average, low relevance, while the other 16 (84%) had high relevance. There were no factors relevant to acceptance and irrelevant to trust (upper-left section in the plot) and vice versa (bottom-right section in the plot).

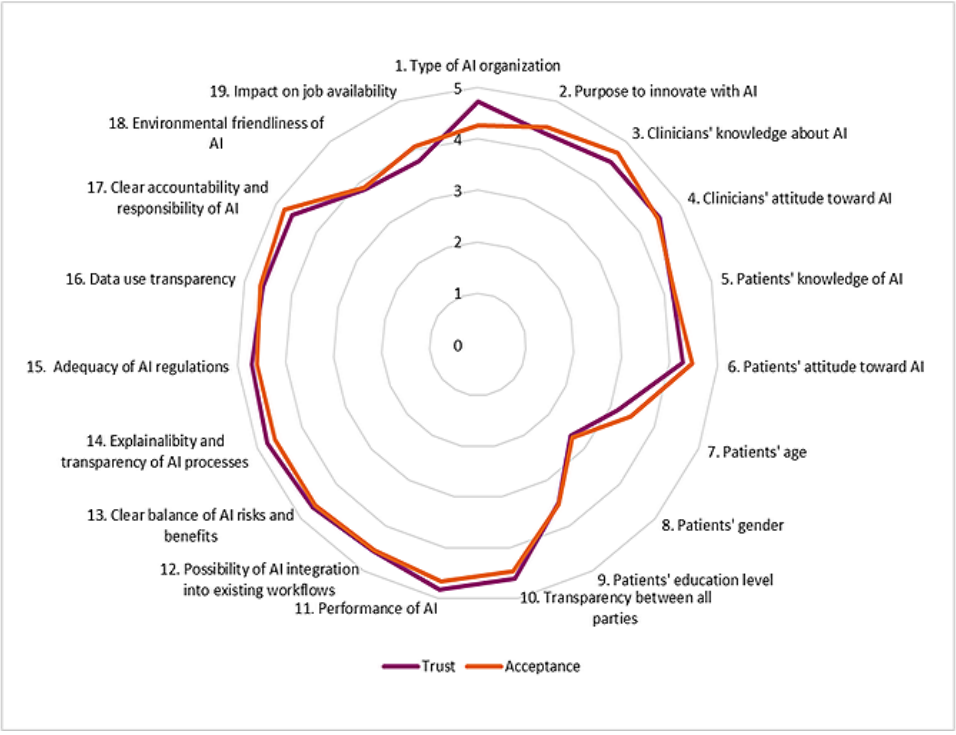
XSL•FO
RenderX

Table 6. Mean (SD) factor relevance to trust and acceptance (N=22).

Factor	Trust	Acceptance
Type of AI[a] organization	4.72 (0.75)	4.27 (0.88)
Purpose to innovate with AI	4.33 (0.84)	4.47 (0.64)
Clinicians' knowledge about AI	4.50 (0.51)	4.73 (0.46)
Clinicians' attitude towards AI	4.50 (0.51)	4.47 (0.64)
Patients' knowledge of AI	4.17 (0.62)	4.20 (0.68)
Patients' attitude toward AI	4.28 (0.57)	4.47 (0.64)
Patients' age	3.17 (1.04)	3.47 (1.19)
Patients' gender	2.61 (1.14)	2.67 (1.05)
Patients' education level	3.50 (0.99)	3.53 (1.19)
Transparency between all parties	4.61 (0.50)	4.47 (0.64)
Performance of AI	4.83 (0.38)	4.67 (0.62)
Possibility of AI integration into existing workflows	4.56 (0.62)	4.53 (0.83)
Clear balance of AI risks and benefits	4.67 (0.49)	4.60 (0.63)
Explainability and transparency of AI processes	4.78 (0.43)	4.60 (0.63)
Adequacy of AI regulations	4.72 (0.57)	4.60 (0.83)
Data use transparency	4.61 (0.50)	4.67 (0.49)
Clear accountability and responsibility of AI	4.61 (0.61)	4.80 (0.41)
Environmental friendliness of AI	3.83 (0.79)	3.87 (0.92)
Impact on job availability	3.78 (1.11)	4.07 (0.88)

[a]AI: artificial intelligence.

Figure 4. Mean scores of factors' relevance to trust in and acceptance of AI applications in medicine (N=19). Score=1 means irrelevant; score=3 means not irrelevant, nor relevant; and score=5 means relevant. AI: artificial intelligence.



XSL•FO
RenderX

Figure 5. Overview of the relevance of factors related to trust in and acceptance of novel AI applications in medicine (1=not relevant, 5=relevant). AI: artificial intelligence.

Factors of Low Relevance

With regard to patients informed about AI application usage in the hospital, participants deemed the patient's gender, age, and educational level to be of low relevance to trust in and acceptance of novel AI applications in medicine.

Factors of High Relevance

The majority of factors were deemed highly relevant to trust in and acceptance of novel AI applications in medicine by participants. Regarding AI professionals, it was observed that the type of institution or organization where AI professionals originated from (eg, university, technology company, commercial organization) and the purpose to innovate with a specific AI application in medicine (eg, financial, societal, or clinical purpose) were considered relevant. Participants reported that the involvement of health care professionals having knowledge of the AI application (eg, by means of training and education) is highly relevant to trust and acceptance. The health care professionals' attitude toward AI application usage in medicine, comprising their agreeableness, openness, conscientiousness, and engagement, was found to be equally important. Likewise, the patients' general knowledge of and attitude toward AI application usage in medicine were found to be relevant. The transparency between all involved parties (AI professionals, health care professionals, and patients) was also deemed highly relevant. Technology-related factors were found to be highly relevant, too, in particular the performance of AI applications in medicine (eg, reproducibility and accuracy of outcomes), the possibility of integration of the AI applications into existing clinical workflows, having a clear balance of risks and benefits of the AI applications, and the explainability and transparency of the processes and outcomes. Legal and ethical factors were also considered of high relevance and concerned the adequacy of the regulations and governance of AI applications in medicine, data use transparency, and clear accountability and responsibility of the AI applications (machine vs human responsibility). Additional factors, such as AI applications being environment friendly and the impact of medical AI on job availability (eg, machines replacing human beings), were viewed as factors of high relevance.

Other Factors

Participants were able to share other factors that were not mentioned in the survey questions. Factors related to trust included solidarity and understanding the bias and interdomain knowledge of AI in software development, data science, and medicine. Other factors related to acceptance were the extent to which alternatives to AI applications are available, the length of experience, transparency about limitations, reproducibility, risks evaluation, resources, and the fear to use an AI application (ie, fear of making the wrong decision or fear of losing control).

Discussion

Principal Findings

This study aimed to identify factors related to trust and acceptance toward medical AI applications by means of a rapid

XSL•FO
RenderX

review and to assess their relevance by conducting a survey. Through the rapid review, 19 key factors related to trust in and acceptance of AI-powered medical technologies were identified and subsequently grouped into 4 categories. Our survey results highlight that of all examined factors, 84% (16/19) were considered highly relevant to trust in and acceptance of novel AI applications in medicine. Only the patient's gender, age, and education level (3/19, 16%) were deemed to be of low relevance by participants.

Comparison With Prior Work

Previous studies have reported that trust in technology is mainly determined by human characteristics [54], technology-related factors [55], and environment-related factors [56], which is in line with the findings of our survey. According to Tran et al [57], who investigated patients' perceived benefits and risks of using digital and AI technology in health care, the important factors to consider are the new technologies requiring an overhaul of the current health care system as human care is being replaced by machines and health care professionals becoming sufficiently equipped with increasing knowledge of AI technology. This highlights the importance of several survey factors, including the possibility of AI integration into existing clinical workflows. Therefore, setting features such as understandability, usability, and user-friendliness (factors that frequently appeared in the rapid review) by AI professionals as key goals in the development of novel AI applications would increase the chances of successful integration of AI technology into health care systems. Tran et al [57] also highlighted the increasing importance of data use transparency toward patients and the acute need for clear accountability and responsibility (machine vs human responsibility) concerning the new technology, which also goes hand in hand with the findings from the rapid review and the survey [57]. The patient data handling must be organized in accordance with the existing data protection regulations in respective countries, with additional precautionary measures due to the sensitive nature of such medical data [57]. Shin et al [58] demonstrated that explainability of AI plays a big role in user trust and attitude toward AI. Explainability, along with transparency, was also found to be highly relevant in our study, especially in relation to the AI application processes and outcomes. In addition, Vourgidis et al [59] recommended that AI systems be regularly checked for being up to date, since today's technology is continuously evolving. This again highlights the relevance of the education of health care professionals, since they are the primary users of medical AI technology and hence need to follow the developments in the field. Yang et al [49] found that gender is not relevant to trust in AI technology in medicine. This agrees with our finding that a patient's gender has low relevance to trust in and acceptance of AI technology in medicine. Contrary to our findings, it has been reported that younger generations in general have more trust and are more likely to accept AI systems compared to older generations [11]. In our survey, the majority of participants were aged 40-60 years and above and they exhibited a solid awareness of and a positive attitude toward AI technology. Gillespie et al [11] also stated that highly educated people (university level) are more likely to trust and accept AI systems compared to those without a university degree. However, our survey showed that a patient's educational level has low relevance to trust in and acceptance of medical AI applications.

Strengths and Limitations

To the best of our knowledge, this is the first study to use a rapid review of the latest literature to identify factors related to trust in and acceptance of AI applications in medicine in order to create a survey to evaluate their relevance and the attitudes of health care stakeholders toward implementation of medical AI applications. However, the study has several limitations. Since a large number of papers and reports in the rapid review did not provide sufficient context for the factors for trust or acceptance, there could have been an increased risk of personal bias during interpretation and categorization of those factors. Furthermore, some studies did not clarify whether the reported factors were related to only trust or only acceptance, which could also lead to possible misinterpretation. To minimize the effect of such bias and misinterpretation, a third reviewer (author HJMV) was consulted in such cases. Another limitation is the relatively small number of papers included in the rapid review, given the breadth of the topic. However, this rapid review was intentionally conducted focusing on the most relevant and recent literature to provide an initial overview and highlight key themes in a time-efficient manner. We aimed to provide a starting point that formed the basis for the survey. In addition, the number of participants included in the survey can be considered relatively low, which was caused by difficulties in recruiting participants and the time-constrained nature of the study. However, sufficient diversity in participant characteristics (ie, gender, age, country of work, and years of professional experience) was achieved, which could be considered more important in terms of validity of the study findings. Even though the survey benefited from a sample with a wide diversity in participant characteristics, one of the limitations to consider is the underrepresentation of certain stakeholder groups, in particular technology providers, policy makers, and hospital staff members other than clinicians. If these groups had been included in the survey, different patterns in factor relevance might have been observed, potentially shedding light on additional concerns or challenges associated with AI applications in medicine. Moreover, when considering the relevance of factors assessed through the survey, which were predominantly highlighted by researchers, it is important to note that these factors might be readily attainable or already well established within this specific stakeholder group. As a result, these factors may not necessarily represent challenges or barriers for this particular group, as they are already well versed in the aspects related to trust and acceptance.

Recommendations for Future Research

The results of this study can be valuable for various stakeholders involved in the implementation of novel AI applications, since trust and acceptance building remains a focus point throughout the different stages, including the pilot, implementation, evaluation, and monitoring phases of the process. In the survey, participants shared other factors related to trust in and acceptance of AI applications in medicine that were not included in the survey. However, due to a lack of context, it is not entirely clear what was meant by some of these factors; since these are

XSL•FO
RenderX

open to interpretation, follow-up research is required to better understand this. In addition, further research is needed to gain insight into the reasons participants considered factors to be of low or high relevance. Regarding the currently underrepresented stakeholder groups in the survey, more research is required to gain insight into the perspectives of policy regulators, technology providers, and hospital staff members. Next, once the implementation of a novel AI technology, such as the AIDPATH system, becomes clear from the trust and acceptance point of view, it would be beneficial to conduct a workshop with experts from the AI and biotechnology fields to identify technical challenges of implementation. This is crucial since, according to the survey results, the technical robustness and clarity of AI applications is a prerequisite for trust and acceptance exhibited toward this technology by stakeholders.

Recommendations for Implementation

By considering the factors that are most relevant in the AI technology adoption process, the implementers can facilitate trust in and acceptance of medical AI applications among their users and other stakeholders. Furthermore, the knowledge of the factors with high relevance to stakeholders can predict concerns the potential users might have regarding the new AI technology and act upon these concerns to implement the AI application efficiently and in a timely manner. There are several ways in how the results of the survey could be used by AI implementers, such as smart hospitals, to build trust and acceptance among various stakeholder groups. For instance, the highly relevant factor of knowledge and understanding of AI among health care professionals could be addressed by providing information about medical AI to clinicians in the form of conferences and educational workshops. These initiatives can ensure that health care professionals remain updated on significant changes in AI technology, facilitating its accurate utilization. Similarly, patients could be informed of medical AI technology through patient information initiatives in (smart) hospitals and within patient communities. The highly relevant technology-related factors could be used by technology developers and scientific researchers as guidance in the development of novel AI technology. For regulators and policy makers, it is crucial to know that users and other stakeholders consider data use transparency and fairness and equity to be of utmost importance regarding novel medical AI technology. Indeed, data privacy is a crucial and ever-so-present topic in legislation and regulations, but it needs to be constantly reviewed by policy makers due to the newness of AI in health care and the speed of its development. The legal aspects of software containing AI have been subjected to the Medical Device Regulation (MDR) [60]. For the acceptance of AI, its implementation in MDR-compliant solutions is invaluable. The tasks of policy makers could involve the risk assessment of various data breaches related to AI in medicine with continuous updating of regulations related to data security and privacy within the field of medical AI. Furthermore, both policy makers and AI professionals have to ensure the maintenance of fairness and equity of AI technology usage.

Conclusion

This study identified and assessed the relevance of factors for trust in and acceptance of AI applications in medicine. The survey demonstrated that the majority of the identified human-related, technology-related, and legal and ethical factors for trust in and acceptance of novel AI applications in medicine were considered by stakeholders to be of high relevance. Taken together, these findings and subsequent recommendations could be used by any implementers of medical AI, such as (smart) hospitals, AI technology organizations, biotechnology research institutes, and policy makers, to facilitate smooth and timely adoption of novel AI applications in medicine.

Acknowledgments

Our gratitude goes out to the AIDPATH (artificial intelligence [AI]-driven Decentralized Production for Advanced Therapies in the Hospital) Consortium partners, the Fraunhofer Institute for Cell Therapy and Immunology IZI, University College London, the Foundation for Research and Technology (FORTH)-Hellas, SZTAKI, Aglaris Cell SL, Sartorius Cell Genix GmbH, the Fundació Clínic per a la Recerca Biomèdica, IRIS Technology Solutions, Red Alert Labs, and ORTEC b.v. We would also like to thank the advisory board members of AIDPATH for providing feedback on the survey. Lastly, we thank the survey's responders.

The paper was written within the framework of the European Union's (EU) Horizon 2020 research and innovation program project AIDPATH (grant agreement number 101016909). All mentioned colleagues/companies are part of AIDPATH and therefore received funding from the EU within the scope of AIDPATH.

Data Availability

The data used and analyzed in this study are available from the corresponding author upon reasonable request.

Authors' Contributions

DS was responsible for methods design, search strategy design, data acquisition, data extraction, data analysis, interpretation of data, and design and writing of the manuscript; AA, methods design, search strategy design, data acquisition, data extraction, data analysis, interpretation of data, regularly reviewing the work, design and writing of the manuscript, and providing feedback on the manuscript; IWAB, methods design, data acquisition, and providing feedback on the manuscript; CS, methods design, regularly reviewing the work, and providing feedback on the manuscript; MH, providing feedback on the manuscript and funding acquisition; JJLJ, reviewing the work and providing feedback on the manuscript; SH, contributing to the Introduction chapter,

XSL•FO
RenderX

providing feedback on the manuscript, project administration, and funding acquisition; and HJMV, concept and design of the overall study, quality assessment, interpretation of data, regularly reviewing the work, providing feedback on the manuscript, and manuscript final approval.

Conflicts of Interest

MH reports speaker honoraria from Novartis, Janssen, and Celgene/BMS and has participated in scientific advisory boards for Janssen and Celgene/BMS. MH is also listed as an inventor on patent applications and has been granted patents related to chimeric antigen receptor (CAR) technologies and CAR T cell therapy that have been filed by the Fred Hutchinson Cancer Research Center and the University of Wurzburg and that have been, in part, licensed to industry. In addition, MH is a cofounder and equity owner of T-CURX GmbH, Wurzburg, Germany. The remaining authors declare that they have no competing interests.

Multimedia Appendix 1
Identified factors and corresponding studies included in the rapid review.
[DOC File , 32 KB - humanfactors_v11i1e47031_app1.doc]

References

1. Artificial intelligence in medicine regulation. European Medicines Agency. URL: https://www.ema.europa.eu/en/news/artificial-intelligence-medicine-regulation [accessed 2022-11-25]
2. Lekadir K, Quaglio G, Tselioudis Garmendia A, Gallin C. Artificial intelligence in healthcare: applications, risks, and ethical and societal impacts. European Parliament. URL: https://www.europarl.europa.eu/stoa/en/document/EPRS_STU(2022)729512 [accessed 2023-12-04]
3. Bohr A, Memarzadeh K. The rise of artificial intelligence in healthcare applications. Artif Intell Healthc 2020 Jun:26-60. [doi: 10.1016/b978-0-12-818438-7.00002-2]
4. Callahan G. What are the advantages of artificial intelligence? rev. URL: https://www.rev.com/blog/speech-to-text-technology/what-are-the-advantages-of-artificial-intelligence [accessed 2023-12-04]
5. Koski E, Murphy J. AI in Healthcare. Stud Health Technol Inform 2021 Dec 15:284-299. [doi: 10.3233/shti210726]
6. Murphy K, Di Ruggiero E, Upshur R, Willison DJ, Malhotra N, Cai JC, et al. Artificial intelligence for good health: a scoping review of the ethics literature. BMC Med Ethics 2021 Feb 15;22(1):14 [FREE Full text] [doi: 10.1186/s12910-021-00577-8] [Medline: 33588803]
7. Rajpurkar P, Chen E, Banerjee O, Topol EJ. AI in health and medicine. Nat Med 2022 Jan 20;28(1):31-38. [doi: 10.1038/s41591-021-01614-0] [Medline: 35058619]
8. Obermeyer Z, Powers B, Vogeli C, Mullainathan S. Dissecting racial bias in an algorithm used to manage the health of populations. Science 2019 Oct 25;366(6464):447-453. [doi: 10.1126/science.aax2342] [Medline: 31649194]
9. Fujimori R, Liu K, Soeno S, Naraba H, Ogura K, Hara K, et al. Acceptance, barriers, and facilitators to implementing artificial intelligence-based decision support systems in emergency departments: quantitative and qualitative evaluation. JMIR Form Res 2022 Jun 13;6(6):e36501 [FREE Full text] [doi: 10.2196/36501] [Medline: 35699995]
10. Schwartz JM, George M, Rossetti SC, Dykes PC, Minshall SR, Lucas E, et al. Factors influencing clinician trust in predictive clinical decision support systems for in-hospital deterioration: qualitative descriptive study. JMIR Hum Factors 2022 May 12;9(2):e33960 [FREE Full text] [doi: 10.2196/33960] [Medline: 35550304]
11. Gillespie N, Lockey S, Curtis C. Trust in artificial intelligence: a five country study. KPMG. 2021. URL: https://assets.kpmg.com/content/dam/kpmg/au/pdf/2021/trust-in-ai-multiple-countries.pdf [accessed 2023-12-04]
12. Ehsan Toreini MA, Coopamootoo K, Elliott K, Gonzalez Zelaya C, van Moorsel A. The relationship between trust in AI and trustworthy machine learning technologies. 2020 Presented at: ACM FAT* 2020: 2020 Conference on Fairness, Accountability, and Transparency; January 27-30, 2020; Barcelona p. 272-283. [doi: 10.1145/3351095.3372834]
13. Ferrario A, Loi M, Viganò E. In AI we trust incrementally: a multi-layer model of trust to analyze human-artificial intelligence interactions. Philos Technol 2019 Oct 23;33(3):523-539. [doi: 10.1007/s13347-019-00378-3]
14. Gursoy D, Chi O, Lu L, Nunkoo R. Consumers acceptance of artificially intelligent (AI) device use in service delivery. Int J Inf Manag 2019 Dec;49:157-169 [FREE Full text] [doi: 10.1016/j.ijinfomgt.2019.03.008]
15. Cornelissen L, Egher C, van Beek V, Williamson L, Hommes D. The drivers of acceptance of artificial intelligence-powered care pathways among medical professionals: web-based survey study. JMIR Form Res 2022 Jun 21;6(6):e33368 [FREE Full text] [doi: 10.2196/33368] [Medline: 35727614]
16. Joshi M, Mecklai K, Rozenblum R, Samal L. Implementation approaches and barriers for rule-based and machine learning-based sepsis risk prediction tools: a qualitative study. JAMIA open. JAMIA Open 2022 Jul;5(2):ooac022. [doi: 10.1093/jamiaopen/ooac022]
17. Rainie L, Anderson J. The fate of online trust in the next decade: theme 3: trust will not grow, but technology usage will continue to rise as a 'new normal' sets in. Pew Research Center. 2023 Jun 1. URL: https://www.pewresearch.org/internet/2017/08/10/theme-3-trust-will-not-grow-but-technology-usage-will-continue-to-rise-as-a-new-normal-sets-in/ [accessed 2023-12-04]

XSL•FO
RenderX

18. Kerasidou CX, Kerasidou A, Buscher M, Wilkinson S. Before and beyond trust: reliance in medical AI. J Med Ethics 2022 Nov 23;48(11):852-856 [FREE Full text] [doi: 10.1136/medethics-2020-107095] [Medline: 34426519]
19. Asan O, Bayrak AE, Choudhury A. Artificial intelligence and human trust in healthcare: focus on clinicians. J Med Internet Res 2020 Jun 19;22(6):e15154 [FREE Full text] [doi: 10.2196/15154] [Medline: 32558657]
20. Eltawil FA, Atalla M, Boulos E, Amirabadi A, Tyrrell PN. Analyzing barriers and enablers for the acceptance of artificial intelligence innovations into radiology practice: a scoping review. Tomography 2023 Jul 28;9(4):1443-1455 [FREE Full text] [doi: 10.3390/tomography9040115] [Medline: 37624108]
21. Matheny M, Israeli S, Ahmed M, Whicher D. Artificial Intelligence in Health Care: The Hope, the Hype, the Promise, the Peril. Washington, DC: The National Academies Press; 2019.
22. AIDPATH. Sciencrew. URL: https://www.sciencrew.com/c/6499?title=AIDPATH [accessed 2023-12-04]
23. Artificial intelligence-driven, decentralized production for advanced therapies in the hospital. European Commission. URL: https://cordis.europa.eu/project/id/101016909 [accessed 2023-01-17]
24. Ran T, Eichmüller SB, Schmidt P, Schlander M. Cost of decentralized CAR T-cell production in an academic nonprofit setting. Int J Cancer 2020 Dec 15;147(12):3438-3445 [FREE Full text] [doi: 10.1002/ijc.33156] [Medline: 32535920]
25. Hoogendoorn K, Main H. What are ATMPs? Claryifying misconceptions. ATMP Sweden. 2020 Jun 12. URL: https://atmpsweden.se/wp-content/uploads/2020/06/What-are-ATMPs-170620.pdf [accessed 2023-12-04]
26. Hort S, Herbst L, Bäckel N, Erkens F, Niessing B, Frye M, et al. Toward rapid, widely available autologous CAR-T cell therapy – artificial intelligence and automation enabling the smart manufacturing hospital. Front Med (Lausanne) 2022 Jun 6;9:913287 [FREE Full text] [doi: 10.3389/fmed.2022.913287] [Medline: 35733863]
27. Arevalo-Rodriguez I, Tricco AC, Steingart KR, Nussbaumer-Streit B, Kaunelis D, Alonso-Coello P, et al. Challenges of rapid reviews for diagnostic test accuracy questions: a protocol for an international survey and expert consultation. Diagn Progn Res 2019 Apr 4;3(1):7 [FREE Full text] [doi: 10.1186/s41512-019-0052-y] [Medline: 31093577]
28. Eysenbach G. Improving the quality of web surveys: the Checklist for Reporting Results of Internet E-Surveys (CHERRIES). J Med Internet Res 2004 Sep 29;6(3):e34 [FREE Full text] [doi: 10.2196/jmir.6.3.e34] [Medline: 15471760]
29. Polland RJ. Essentials of survey research and analysis. Psycholosphere. 1998. URL: https://www.psycholosphere.com/Essentials%20of%20Survey%20Research%20and%20Analysis%20by%20Ronald%20Polland.pdf [accessed 2023-12-04]
30. 30. Act on medical-scientific research with humans. Central Committee on Research Involving Human Subjects. URL: https://wetten.overheid.nl/BWBR0009408/2022-07-01 [accessed 2023-09-11]
31. Alaqra AS, Kane B, Fischer-Hübner S. Machine learning-based analysis of encrypted medical data in the cloud: qualitative study of expert stakeholders' perspectives. JMIR Hum Factors 2021 Sep 16;8(3):e21810 [FREE Full text] [doi: 10.2196/21810] [Medline: 34528892]
32. Allen D. Trust and artificial intelligence in healthcare. MedicalExpo e-Magazine. URL: https://emag.medicalexpo.com/trust-and-artificial-intelligence-in-healthcare/ [accessed 2023-12-04]
33. Antes AL, Burrous S, Sisk BA, Schuelke MJ, Keune JD, DuBois JM. Exploring perceptions of healthcare technologies enabled by artificial intelligence: an online, scenario-based survey. BMC Med Inform Decis Mak 2021 Jul 20;21(1):221 [FREE Full text] [doi: 10.1186/s12911-021-01586-8] [Medline: 34284756]
34. Brady D. Building public trust to successfully deploy AI in EU healthcare. European Policy Centre. 2022 Jan 26. URL: https://www.epc.eu/content/PDF/2022/CHES_PB_v2-2.pdf [accessed 2023-12-04]
35. Diprose WK, Buist N, Hua N, Thurier Q, Shand G, Robinson R. Physician understanding, explainability, and trust in a hypothetical machine learning risk calculator. J Am Med Inform Assoc 2020 Apr 01;27(4):592-600 [FREE Full text] [doi: 10.1093/jamia/ocz229] [Medline: 32106285]
36. Gille F, Jobin A, Ienca M. What we talk about when we talk about trust: theory of trust for AI in healthcare. Intell-Based Med 2020 Nov;1-2:100001-100002 [FREE Full text] [doi: 10.1016/j.ibmed.2020.100001]
37. Haggenmüller S, Krieghoff-Henning E, Jutzi T, Trapp N, Kiehl L, Utikal JS, et al. Digital natives' preferences on mobile artificial intelligence apps for skin cancer diagnostics: survey study. JMIR Mhealth Uhealth 2021 Aug 27;9(8):e22909 [FREE Full text] [doi: 10.2196/22909] [Medline: 34448722]
38. Hui CY, McKinstry B, Fulton O, Buchner M, Pinnock H. Patients' and clinicians' perceived trust in internet-of-things systems to support asthma self-management: qualitative interview study. JMIR Mhealth Uhealth 2021 Jul 16;9(7):e24127 [FREE Full text] [doi: 10.2196/24127] [Medline: 34269684]
39. Jauk S, Kramer D, Avian A, Berghold A, Leodolter W, Schulz S. Correction to: technology acceptance of a machine learning algorithm predicting delirium in a clinical setting: a mixed-methods study. J Med Syst 2021 Mar 10;45(4):52 [FREE Full text] [doi: 10.1007/s10916-021-01728-5] [Medline: 33740133]
40. Jermutus E, Kneale D, Thomas J, Michie S. Influences on user trust in healthcare artificial intelligence: a systematic review. Wellcome Open Res 2022 Feb 18;7:65. [doi: 10.12688/wellcomeopenres.17550.1]
41. Jobin A, Ienca M, Vayena E. The global landscape of AI ethics guidelines. Nat Mach Intell 2019 Sep 02;1(9):389-399. [doi: 10.1038/s42256-019-0088-2]
42. Kovalchuk S, Kopanitsa G, Derevitskii I, Matveev G, Savitskaya D. Three-stage intelligent support of clinical decision making for higher trust, validity, and explainability. J Biomed Inform 2022 Mar;127:104013 [FREE Full text] [doi: 10.1016/j.jbi.2022.104013] [Medline: 35158071]

XSL•FO
RenderX

43. Shah P, Mishra D, Shanmugam M, Vighnesh M, Jayaraj H. Acceptability of artificial intelligence-based retina screening in general population. Indian J Ophthalmol 2022;70(4):1140. [doi: 10.4103/ijo.ijo_1840_21]
44. Tamori H, Yamashina H, Mukai M, Morii Y, Suzuki T, Ogasawara K. Acceptance of the use of artificial intelligence in medicine among Japan's doctors and the public: a questionnaire survey. JMIR Hum Factors 2022 Mar 16;9(1):e24680. [doi: 10.2196/24680]
45. Tucci V, Saary J, Doyle TE. Factors influencing trust in medical artificial intelligence for healthcare professionals: a narrative review. J Med Artif Intell 2022 Mar;5:4-4. [doi: 10.21037/jmai-21-25]
46. van Bussel MJP, Odekerken-Schröder GJ, Ou C, Swart RR, Jacobs MJG. Analyzing the determinants to accept a virtual assistant and use cases among cancer patients: a mixed methods study. BMC Health Serv Res 2022 Jul 09;22(1):890 [FREE Full text] [doi: 10.1186/s12913-022-08189-7] [Medline: 35804356]
47. Walter S, Gruss S, Frisch S, Liter J, Jerg-Bretzke L, Zujalovic B, et al. "What about automated pain recognition for routine clinical use?" A survey of physicians and nursing staff on expectations, requirements, and acceptance. Front Med (Lausanne) 2020 Dec 21;7:566278 [FREE Full text] [doi: 10.3389/fmed.2020.566278] [Medline: 33409286]
48. Williams J, Sim A, Lee B, Colon A, Lee R, Schick A, et al. Trustworthy AI (TAI) playbook. U.S. DEPARTMENT OF HEALTH & HUMAN SERVICES. 2021 Sep. URL: https://www.hhs.gov/sites/default/files/hhs-trustworthy-ai-playbook.pdf [accessed 2023-12-04]
49. Yang K, Zeng Z, Peng H, Jiang Y. Attitudes of Chinese cancer patients toward the clinical use of artificial intelligence. Patient Prefer Adher 2019 Nov;13:1867-1875. [doi: 10.2147/ppa.s225952]
50. Yap A, Wilkinson B, Chen E. Patients perceptions of artificial intelligence in diabetic eye screening. Asia-Pac J Ophthalmol (Philadelphia, Pa) 2022 May 1;11(3):287-293. [doi: 10.1097/apo.0000000000000525]
51. Zhang Z, Citardi D, Wang D, Genc Y, Shan J, Fan X. Patients' perceptions of using artificial intelligence (AI)-based technology to comprehend radiology imaging data. Health Inform J 2021 Apr 29;27(2):14604582211011215 [FREE Full text] [doi: 10.1177/14604582211011215] [Medline: 33913359]
52. Zhang Z, Genc Y, Wang D, Ahsen ME, Fan X. Effect of AI explanations on human perceptions of patient-facing AI-powered healthcare systems. J Med Syst 2021 May 04;45(6):64. [doi: 10.1007/s10916-021-01743-6] [Medline: 33948743]
53. Zheng B, Wu M, Zhu S, Zhou H, Hao X, Fei F, et al. Attitudes of medical workers in China toward artificial intelligence in ophthalmology: a comparative survey. BMC Health Serv Res 2021 Oct 09;21(1):1067 [FREE Full text] [doi: 10.1186/s12913-021-07044-5] [Medline: 34627239]
54. Hengstler M, Enkel E, Duelli S. Applied artificial intelligence and trust—the case of autonomous vehicles and medical assistance devices. Technol Forecast Soc Change 2016 Apr;105:105-120 [FREE Full text] [doi: 10.1016/j.techfore.2015.12.014]
55. Schaefer KE, Chen JYC, Szalma JL, Hancock PA. A meta-analysis of factors influencing the development of trust in automation: implications for understanding autonomy in future systems. Hum Factors 2016 May 22;58(3):377-400. [doi: 10.1177/0018720816634228] [Medline: 27005902]
56. Oleson K, Billings D, Kocsis V, Chen J, Hancock P. Antecedents of trust in human-robot collaborations. 2011 Presented at: 2011 IEEE International Multi-Disciplinary Conference on Cognitive Methods in Situation Awareness and Decision Support (CogSIMA); February 22-24, 2011; Miami Beach, FL p. 175-178. [doi: 10.1109/cogsima.2011.5753439]
57. Tran V, Riveros C, Ravaud P. Patients' views of wearable devices and AI in healthcare: findings from the ComPaRe e-cohort. NPJ Digit Med 2019 Jun 14;2(1):53 [FREE Full text] [doi: 10.1038/s41746-019-0132-y] [Medline: 31304399]
58. Shin D. The effects of explainability and causability on perception, trust, and acceptance: implications for explainable AI. Int J Hum-Comput Stud 2021 Feb;146:102551 [FREE Full text] [doi: 10.1016/j.ijhcs.2020.102551]
59. Vourgidis I, Mafuma S, Wilson P, Carter J, Cosma G. Medical Expert Systems – A Study of Trust and Acceptance by Healthcare Stakeholders. Berlin: Springer International Publishing; 2019:108-119.
60. EUR-Lex. Regulation (EU) 2017/745 of the European Parliament and of the Council of 5 April 2017 on medical devices, amending Directive 2001/83/EC, Regulation (EC) No 178/2002 and Regulation (EC) No 1223/2009 and repealing Council Directives 90/385/EEC and 93/42/EEC. European Parliament. 2017. URL: https://eur-lex.europa.eu/eli/reg/2017/745/oj [accessed 2023-12-04]

Abbreviations

AI: artificial intelligence
AIDPATH: AI-powered Decentralized Production for Advanced Therapies in the Hospital
CAR-T: chimeric antigen receptor T cell
MDR: Medical Device Regulation

XSL•FO
RenderX

Edited by A Kushniruk, E Borycki; submitted 06.03.23; peer-reviewed by D Kohen, Y Zhuang, J Zeng; comments to author 17.06.23; revised version received 25.09.23; accepted 20.11.23; published 17.01.24.

Please cite as:
Shevtsova D, Ahmed A, Boot IWA, Sanges C, Hudecek M, Jacobs JJL, Hort S, Vrijhoef HJM
Trust in and Acceptance of Artificial Intelligence Applications in Medicine: Mixed Methods Study
JMIR Hum Factors 2024;11:e47031
URL: https://humanfactors.jmir.org/2024/1/e47031
doi:10.2196/47031
PMID:38231544

©Daria Shevtsova, Anam Ahmed, Iris W A Boot, Carmen Sanges, Michael Hudecek, John J L Jacobs, Simon Hort, Hubertus J M Vrijhoef. Originally published in JMIR Human Factors (https://humanfactors.jmir.org), 17.01.2024. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on https://humanfactors.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

A Machine Learning Approach with Human-AI Collaboration for Automated Classification of Patient Safety Event Reports: Algorithm Development and Validation Study

Hongbo Chen[1], MSc; Eldan Cohen[1], PhD; Dulaney Wilson[2], PhD; Myrtede Alfred[1], PhD

[1]Department of Mechanical & Industrial Engineering, Faculty of Applied Science & Engineering, University of Toronto, Toronto, ON, Canada
[2]Department of Public Health Sciences, College of Medicine, Medical University of South Carolina, Charleston, SC, United States

Corresponding Author:
Myrtede Alfred, PhD
Department of Mechanical & Industrial Engineering
Faculty of Applied Science & Engineering
University of Toronto
27 King's College Cir
Toronto, ON, M5S 1A1
Canada
Phone: 1 4372154739
Email: myrtede.alfred@utoronto.ca

Abstract

Background: Adverse events refer to incidents with potential or actual harm to patients in hospitals. These events are typically documented through patient safety event (PSE) reports, which consist of detailed narratives providing contextual information on the occurrences. Accurate classification of PSE reports is crucial for patient safety monitoring. However, this process faces challenges due to inconsistencies in classifications and the sheer volume of reports. Recent advancements in text representation, particularly contextual text representation derived from transformer-based language models, offer a promising solution for more precise PSE report classification. Integrating the machine learning (ML) classifier necessitates a balance between human expertise and artificial intelligence (AI). Central to this integration is the concept of explainability, which is crucial for building trust and ensuring effective human-AI collaboration.

Objective: This study aims to investigate the efficacy of ML classifiers trained using contextual text representation in automatically classifying PSE reports. Furthermore, the study presents an interface that integrates the ML classifier with the explainability technique to facilitate human-AI collaboration for PSE report classification.

Methods: This study used a data set of 861 PSE reports from a large academic hospital's maternity units in the Southeastern United States. Various ML classifiers were trained with both static and contextual text representations of PSE reports. The trained ML classifiers were evaluated with multiclass classification metrics and the confusion matrix. The local interpretable model-agnostic explanations (LIME) technique was used to provide the rationale for the ML classifier's predictions. An interface that integrates the ML classifier with the LIME technique was designed for incident reporting systems.

Results: The top-performing classifier using contextual representation was able to obtain an accuracy of 75.4% (95/126) compared to an accuracy of 66.7% (84/126) by the top-performing classifier trained using static text representation. A PSE reporting interface has been designed to facilitate human-AI collaboration in PSE report classification. In this design, the ML classifier recommends the top 2 most probable event types, along with the explanations for the prediction, enabling PSE reporters and patient safety analysts to choose the most suitable one. The LIME technique showed that the classifier occasionally relies on arbitrary words for classification, emphasizing the necessity of human oversight.

Conclusions: This study demonstrates that training ML classifiers with contextual text representations can significantly enhance the accuracy of PSE report classification. The interface designed in this study lays the foundation for human-AI collaboration in the classification of PSE reports. The insights gained from this research enhance the decision-making process in PSE report classification, enabling hospitals to more efficiently identify potential risks and hazards and enabling patient safety analysts to take timely actions to prevent patient harm.

(JMIR Hum Factors 2024;11:e53378) doi:10.2196/53378

XSL•FO
RenderX

KEYWORDS

accident; accidents; black box; classification; classifier; collaboration; design; document; documentation; documents; explainability; explainable; human-AI collaboration; human-AI; human-computer; human-machine; incident reporting; interface design; interface; interpretable; LIME; machine learning; patient safety; predict; prediction; predictions; predictive; report; reporting; safety; text; texts; textual; artificial intelligence

Introduction

Since the publication of the seminal report on patient safety—*To Err Is Human* [1], the importance of preventing adverse events in health care has been widely recognized. Adverse events refer to unintended or unexpected incidents that occur during hospital care that cause harm to a patient [2]. Common adverse events include complications, falls, and medication errors. These events can lead to prolonged hospital stays, permanent harm to patients, life-saving interventions, or even contributing to patient deaths [2,3]. Unfortunately, adverse events remain one of the top 10 leading causes of death and disability worldwide, resulting in 251,454 deaths annually in the United States alone [4]. In Organization for Economic Cooperation and Development (OECD) countries, 15% of total hospital activity is the direct result of adverse events [5]. The global cost of adverse events has been estimated at 42 billion USD annually [6].

Patient safety event (PSE) reporting systems, also called incident reporting systems, have been widely adopted in hospitals across the world as part of their efforts to mitigate adverse events and improve patient safety [7,8]. Multiple nations, including Canada, Japan, England, and Norway, have made it mandatory for hospitals to establish and maintain a PSE reporting system, either with individual health care systems or through centralized national incident reporting platforms [9]. The primary purpose of the PSE reporting system is to provide health care organizations with a centralized system for tracking and analyzing PSEs, thereby facilitating continuous learning and maintaining a record of PSEs for risk assessment and prevention [7,10]. PSE reporting systems are tools that allow frontline health care personnel to voluntarily report adverse events, near-misses, and unsafe conditions [11]. Each PSE report includes structured data, such as event types, patient harm level, date, and location of the event, as well as unstructured data, including a free-text section that contains the factual description of the event and the patient's outcome [12]. Following submission, PSE reports are reviewed by relevant hospital staff, such as risk managers, patient safety analysts, nurse managers, physicians, and biomedical engineers, to identify areas for patient safety and quality improvement within the hospital [13].

Accurately classifying PSE reports into their appropriate event type is crucial to ensure that these reports are directed to the relevant patient safety analyst, support organizational learning, identify patterns and trends in adverse events, and ultimately prioritize measures to reduce adverse events [14,15]. An event type refers to a specific class of events that share common characteristics [16]. Examples of event types include falls, medication-related issues, and diagnosis errors [17,18]. PSE reporting systems may have upwards of 20 categories of events. The formulation of these classification taxonomies generally involves systematically grouping PSE reports based on common characteristics [19]. The descriptions of event types are not always readily accessible to PSE reporters and patient safety analysts [15]. Previous studies have found that the classification of PSE reports is inconsistent depending on the reporter's profession, interpretation of the adverse event, and understanding of the PSE classification taxonomy [15,20]. Furthermore, 25% of PSE reports are labeled with vague or nonspecific categories such as "miscellaneous" and "other" and require time-consuming retrospective analysis for reclassification [21]. These problems are further exacerbated by the growing volume of PSEs reported [18,22]. For instance, hospitals in the state of New South Wales in Australia reported close to 195,000 PSEs in 2020 [23], while there were approximately 2.3 million PSEs reported to the National Reporting and Learning System in England from April 2021 to March 2022 [24].

In light of these challenges, it is imperative to find an efficient solution to ensure the reliable classification of PSE reports. Recent studies have used static text representations and supervised machine learning (ML) techniques to automate the PSE report classification [17,25,26]. However, static text representations ignore the ordering of the words and do not account for the differences in word meaning across different contexts. These limitations may result in suboptimal classification performance. With the emergence of deep learning, contextual text representation produced from transformer-based deep learning models has achieved state-of-the-art performance on a wide range of natural language processing tasks, including text classification [27]. The contextual representation of each word is based on its surrounding context within the text, allowing for a more accurate understanding of its usage across different contexts and facilitating knowledge transfer across languages [28]. Therefore, using contextual text representation in training ML classifiers presents a promising opportunity for achieving a more precise classification of PSE reports.

The integration of ML models into PSE reporting systems has important implications for human–artificial intelligence (AI) collaboration, given the roles of the incident reporter (front end) and patient safety analyst (backend). Various approaches for using ML classifiers can be developed, including at different levels of automation; however, unifying the strengths of both human expertise and AI offers the most promising route for effective implementation [29-31]. A crucial determinant for successfully implementing the human-AI collaboration approach is decision transparency [32,33], which is often referred to as explainability. Explainability is the concept that an ML model's prediction can be explained in a way that human operators can comprehend and reconstruct the model's reasoning [33]. Incorporating explainability techniques in human-AI collaboration is paramount as it facilitates a deeper understanding of the factors influencing the predictions, thereby fostering trust and understanding between human experts and AI systems. Therefore, embedding explainability into the

XSL•FO
RenderX

human-AI collaboration holds significant potential for enhancing PSE report classification.

The main aim of this study is to examine the efficacy of contextual text representation in improving the accuracy of PSE report classification. To accomplish this, we trained, evaluated, and compared various ML classifiers with both static and contextual text representations. Additionally, we developed an interface to illustrate the integration of the ML classifier in an event reporting system to support human-AI collaboration for PSE report classification. Moreover, we enhanced the explainability of the ML classifiers by using an explainable AI technique. Furthermore, we have investigated the ML classifier's performance under 2 conditions, differentiated by whether the explanation is valid for the predicted event type. Based on this analysis, we offer recommendations for optimizing human-AI collaboration in the context of PSE report classification.

Methods

Data Collection

The data set for this study was obtained from a large academic hospital located in the Southeastern United States. A total of 861 PSE reports from the labor and delivery and mother-baby units were extracted from the PSE reporting system from January 1, 2019, to December 31, 2020. Each PSE report was assigned to a single event type from a set of 25 classes, such as complication of the surgery, fall, medication-related, and supply issues. The ML classifiers were trained exclusively on PSE reports from the 7 most frequently occurring event types. This selection was intended to create a more balanced training data set to reduce sampling bias and the risk of overfitting. The selected PSE reports used for training ML classifiers constitute approximately 72.8% (627/861) of the extracted reports (Table 1).

Table 1. Prevalence of patient safety event reports by event type in this study.

Event type	Extracted reports (n=861), n (%)
Care coordination or communication	186 (21.6)
Laboratory test	122 (14.2)
Medication related	89 (10.3)
Omission or errors in assessment, diagnosis, and monitoring	67 (7.8)
Maternal	58 (6.7)
Equipment or devices	56 (6.5)
Supplies	49 (5.7)
Total	627 (72.8)

Data Preprocessing

The free-text section of PSE reports was preprocessed before feeding into ML classifiers as input features. The preprocessing procedures include text normalization, feature extraction, data splitting, and data augmentation (Multimedia Appendix 1 [28,34-39]).

Classifier Training

A range of ML classifiers, including multinomial logistic regression (MLR), support vector machine (SVM), extreme gradient boosting, light gradient boosting, random forest (RF), *k*-nearest neighbor (KNN), and multilayer perceptron, were used for the classification of PSE reports. While SVM is a binary classifier, it is also capable of performing multiclass classification using the one-versus-one strategy. This involves treating the multiclass classification problem as a series of binary classification problems, creating $n \times (n-1)/2$ binary classifiers for each pair of classes, where n represents the total number of classes, and the final classification is based on the majority vote of all binary classifiers. Extreme gradient boosting, light gradient boosting, and RF are tree-based ensemble algorithms that are commonly used in text classification tasks [17,40]. The KNN classifier predicts the class of a data point based on the majority class among its nearest neighbors in the training data set. Multilayer perceptron is a feedforward neural network consisting of multiple layers of interconnected neutrons and trained using backpropagation.

To optimize the performance of ML classifiers, we used the 5-fold cross-validation grid search technique to identify the best combination of hyperparameters. During this process, a range of values of important hyperparameters (ie, regularization strength) is assessed with 5-fold cross-validation. For each combination of hyperparameters, the training set is randomly split into 5 distinct folds, and then the ML classifier is trained and evaluated 5 times, picking a different fold for evaluation every time and training on the remaining 4 folds. The optimized combination of hyperparameters is determined based on the average performance of the classifier on the F_1-score across the 5-fold cross-validation runs.

Classifier Evaluation

We evaluated the performance of the trained classifiers on the testing set with standard classification metrics, including accuracy, precision, recall, F_1-score, and area under the receiver operating characteristic curve. We also evaluated classifiers on top-2 accuracy, which measures the proportion of predictions where the correct event type is among the top 2 highest probability event types predicted by the classifier. The definitions and mathematical formulas of the evaluation metrics are shown in Multimedia Appendix 2. Each of these metrics provides a distinct perspective on the performance of the

XSL•FO
RenderX

classifier, and collectively, they offer a comprehensive understanding of how well the classifier is functioning. Since we framed PSE report classification as a multiclass text classification problem, the precision, recall, F_1-score, and area under the receiver operating characteristic curve are computed for each class and combined using a weighted average where the weights correspond to the number of data points in each class.

Development and Assessment of Explainability

As the contextual text representation is generated from transformer-based neural network, which has a black box nature, we used the local interpretable model-agnostic explanations (LIME) technique to analyze the top-performing ML classifier trained with the contextual text representation. LIME is a post hoc, local perturbation technique that provides the explanation for a single prediction. LIME generates perturbed data by randomly removing words from a text document and trains a locally explainable model with perturbed data to simulate the original classifier's prediction [41]. By measuring how the classifier's prediction changes under these perturbations, LIME reflects the contributions of each word to the prediction. The importance of each word can then be assessed for a single prediction, revealing whether the ML classifier has learned to use relevant words for classifying PSE reports. We used LIME to generate explanations for the top-performing classifier's prediction, specifically by highlighting the words that the classifier deems influential for the prediction. We presented 3 distinct cases: one where the classifier effectively leveraged relevant words for accurate prediction, another where it failed to do so, and a final case that illustrated the explanation for a misclassification. In addition, we analyzed the top 5 most prevalent words identified by LIME for each event type.

A total of 2 human factors graduate students were recruited to assess the quality of the LIME explanations. For each PSE report in the test data set, the reviewers were asked to determine independently if any of the highlighted words were relevant to the predicted event type. Based on these evaluations, the reports were then categorized into 2 distinct groups: those in which the highlighted terms were deemed relevant to the predicted event types and those where they were deemed irrelevant. Discrepancies were resolved through discussions. The interrater reliability index (Cohen κ) was calculated to quantify the level of agreement between the reviewers. The ML classifier's accuracy and F_1-score were evaluated for these 2 groups of PSE reports. A subsequent comparison will explore the influence of explanation quality on prediction reliability.

Interface Development

In the typical workflow of PSE report classification, reporters need to provide a narrative description of the event as well as key attributes such as the event type, level of harm, date, and location of the event. Subsequent to this initial classification, the patient safety analyst will review the submitted report and decide if it needs to be recategorized to better reflect the nature of the event [17,42]. To support efficient and reliable categorization, the classifier will need to provide reporters with real-time support during the reporting process. We developed a PSE reporting interface to illustrate the integration of the ML classifier and the LIME explainability technique. In the design, the ML classifier provides multiple high-probability event types along with explanations for its prediction and allows the user to select the most appropriate event type. The interface was developed in Figma [43] and designed using guidance from previous research on incident reporting systems, including question type, mandatory and optional questions, and taxonomy for event type and harm level [44,45].

Ethical Considerations

The study was approved by the Medical University of South Carolina Hospital's institutional review board (Pro00105892). Following data extraction, PSE reports were anonymized in accordance with privacy regulation guidelines.

Results

Performance Comparison

We evaluated the trained ML classifier's classification performance on both static and contextual text representations (Multimedia Appendix 3). The performance of the top-performing ML classifier trained with static and contextual text representations is shown in Table 2. Our results showed that for static text representation, the MLR classifier trained with term frequency–inverse document frequency (TF-IDF) achieved the best performance, with an F_1-score of 0.631 and an accuracy of 66.7% (84/126). On the other hand, for contextual text representation, the SVM classifier trained with RoBERTa-base outperformed others, with an F_1-score of 0.753 and an accuracy of 75.4% (95/126). The SVM classifier trained with RoBERTa-base showed a 19.3% elative improvement in F_1-score and a 13% (11/85) relative improvement in accuracy compared to the MLR classifier trained with TF-IDF for contextual text representation. In addition, we compared the accuracy (95/126, 75.4%) and top 2 accuracy (107/126, 84.9%) of the SVM classifier trained with RoBERTa-base and observed that 9.5% (12/126) of PSE reports' true event type was predicted as the second highest probability event type by the classifier, which represents 39% (12/31) of misclassified PSE reports.

XSL•FO
RenderX

Table 2. Performance of top-performing ML classifiers trained with static and contextual text representations.

Metric	Top-performing ML[a] model trained with the static text representation			Top-performing ML model trained with the contextual text representation		
	Performance	ML classifier	Text Representation	Performance	ML classifier	Text Representation
Accuracy (%)	66.67	MLR[b]	TF-IDF[c]	75.40	SVM[d]	RoBERTa-base
Top 2 accuracy (%)	85.71	MLR	TF-IDF	88.10	MLP[e]	xlm-RoBERTa-base
Precision	0.707	KNN[f]	TF-IDF	0.757	SVM	RoBERTa-base
Recall	0.667	MLR	TF-IDF	0.754	SVM	RoBERTa-base
F_1-score	0.631	MLR	TF-IDF	0.753	SVM	RoBERTa-base

[a]ML: machine learning.
[b]MLR: multinomial logistic regression.
[c]TF-IDF: term frequency–inverse document frequency.
[d]SVM: support vector machine.
[e]MLP: multilayer perceptron.
[f]KNN: *k*-nearest neighbor.

Performance on Classifying Individual Event Types

We analyzed the performance of the SVM classifier trained with RoBERTa-base on individual event types (Table 3). The F_1-score measure for different event types ranged from 0.958 (laboratory test) to 0.400 (omission or errors in assessment, diagnosis, and monitoring).

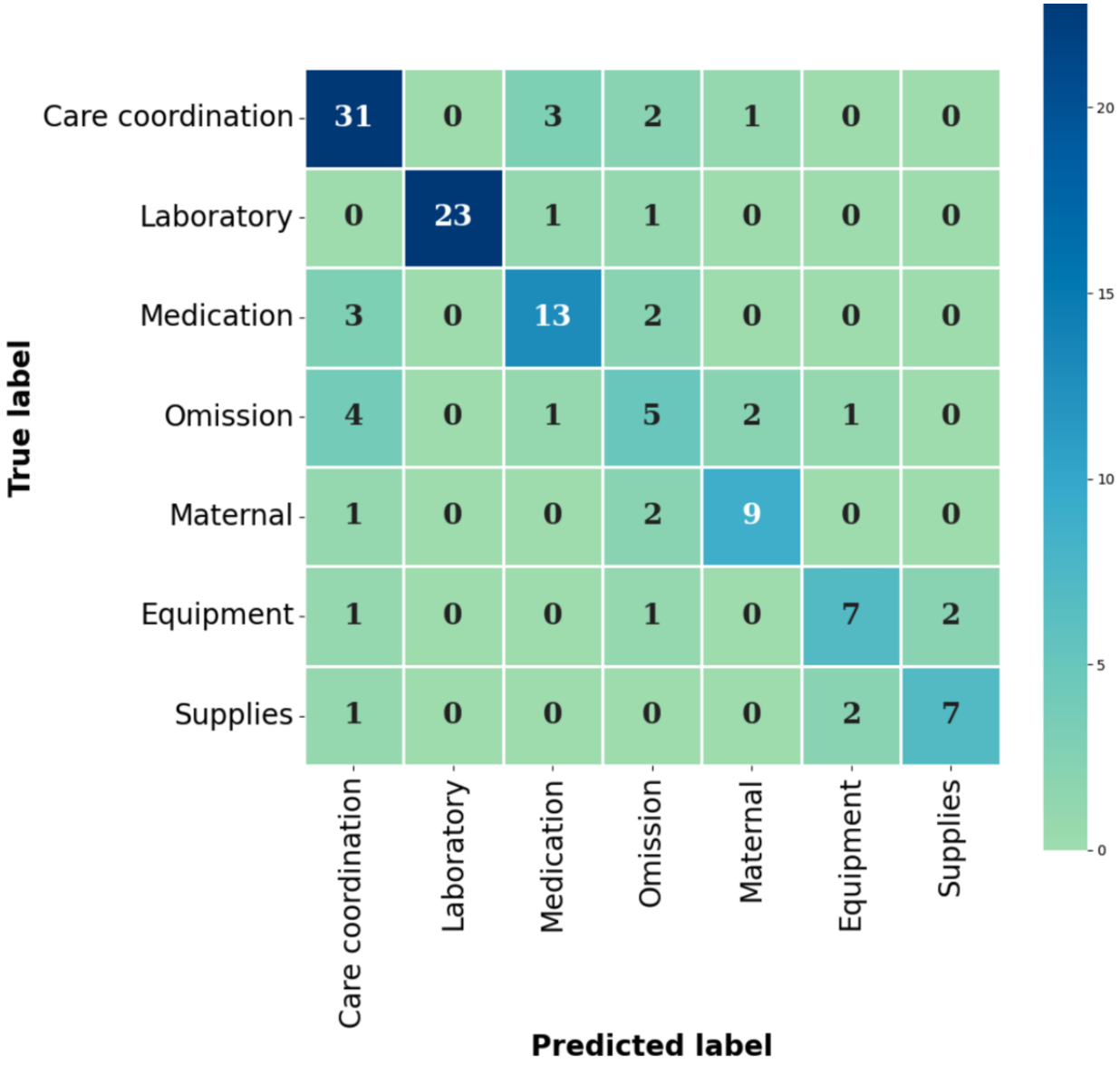
Table 3. Performance of support vector machine+RoBERTa-base on the individual event type.

Event type	Precision	Recall	F_1-score
Care coordination or communication	0.721	0.838	0.775
Laboratory test	1.000	0.920	0.958
Medication related	0.765	0.722	0.743
Omission or errors in assessment, diagnosis, and monitoring	0.417	0.385	0.400
Maternal	0.750	0.750	0.750
Equipment or devices	0.700	0.636	0.667
Supplies	0.778	0.700	0.737

Figure 1 shows the confusion matrix for the SVM classifier trained with RoBERTa-base evaluated on the test set. A confusion matrix is a table that visualizes the performance of a classifier. The main diagonal value is the number of PSE reports that have been classified as true event types, whereas off-diagonal values are the number of PSE reports that have been wrongly classified. While the classifier was able to classify the majority of event types of PSE reports correctly, there is a consistent misclassification of the omission or errors in assessment, diagnosis, or monitoring PSE report as the care coordination or communication (coordination) event type.

XSL•FO
RenderX

Figure 1. Confusion matrix for the testing set evaluation with a support vector machine classifier trained with RoBERTa-base.



LIME-Based Explainability Analysis

We used LIME to evaluate whether the SVM classifier trained with RoBERTa-base has leveraged informative words for classification. Figure 2 presents 3 examples of explanations for the classifier's predictions. At the top of Figure 2, LIME identified "ketorolac," "ibuprofen," and "doses" from the PSE report as important words for classifying the report into the medication-related event type, which is reasonable given the report's association with incorrect medication doses. Conversely, in the middle of Figure 2, LIME highlighted "our," "handle," and "or" from the text as important words for classifying the report into the equipment or device event type. Although the predicted event type was correct, the classifier relied on irrelevant words for the classification. At the bottom of Figure 2, a case of misclassification is shown. LIME highlighted "pitocin," "pump," "available," and "use" as influential words for classifying the PSE report into medication-related event type when it belongs to the equipment class. In addition, for each event type, we extract the 5 most prevalent words that were deemed important for the classifier's prediction across the whole data set (Table 4). This inclusion of stop words (ie, "was," "not," and "till") among influential terms, as shown in Table 4, demonstrated that the classifier does not always rely on relevant words for making classifications.

XSL•FO
RenderX

Figure 2. Local interpretable model-agnostic explanations of support vector machine classifiers trained with RoBERTa-base. MD: medical doctor; PSE: patient safety event; pt: patient.

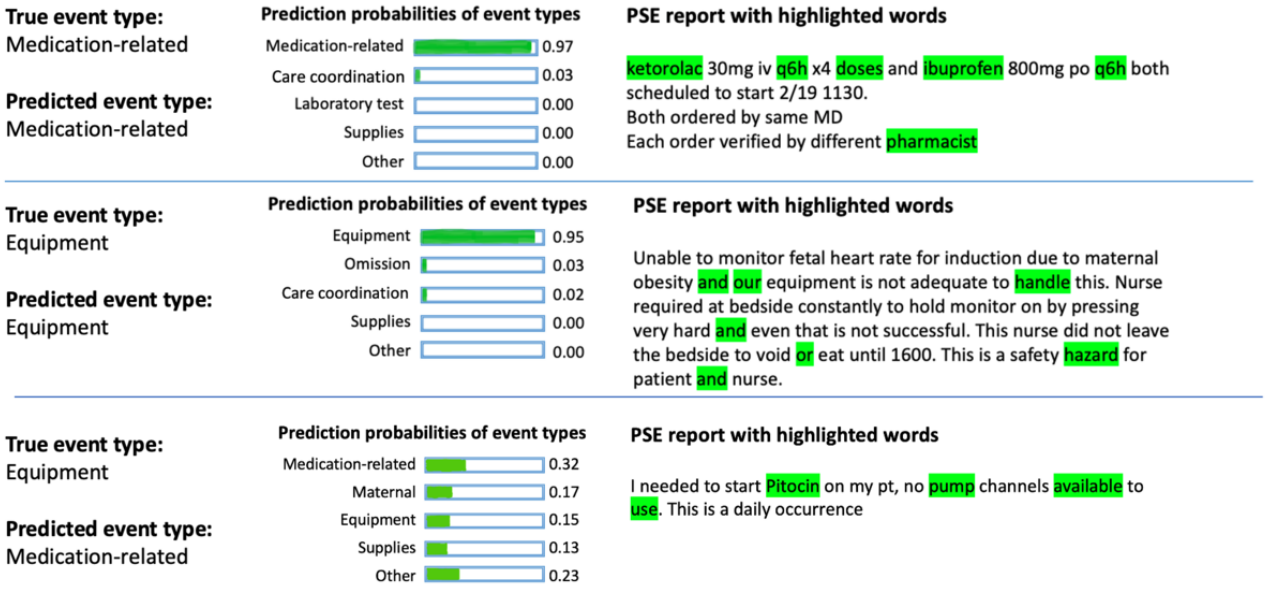


Table 4. The 5 most prevalent and important words for each event type were derived from the support vector machine classifier trained with RoBERTa-base.

Event type	Prevalent influential words highlighted by local interpretable model-agnostic explanations
Care coordination or communication	requested, delayed, patient, not, follow
Laboratory test	specimen, lab, labels, collection, results
Medication related	patches, doses, orders, medication, pitocin
Omission or errors in assessment, diagnosis, and monitoring	warning, patient, was, till, late
Maternal	baby, hysterectomy, stable, pumping, hemorrhage
Equipment or devices	instruments, trays, notified, malfunctioning, faulty
Supplies	vendor, sterile, available, needed, OR

After reviewing the LIME explanations for each PSE report in the test data set, 73.8% (93/126) of the reports were categorized into a subset where at least 1 highlighted word was deemed relevant to the predicted event type. The remaining reports comprised a second subset where no highlighted words were relevant. The interrater reliability index measured by Cohen κ between the 2 reviewers was 0.83, indicating substantial agreement. Table 5 presents the performance of the top-performing ML classifier for both subsets. For the first subset, the classifier achieved an accuracy of 84% (78/93) and an F_1-score of 0.825. In contrast, the second subset showed a classifier accuracy of 52% (17/33) and an F_1-score of 0.549.

Table 5. Performance of a top-performing machine learning classifier on reports that have relevant words highlighted and reports with irrelevant words highlighted.

Metric	PSE[a] reports with relevant words highlighted	PSE reports with irrelevant words highlighted
Number of PSE reports, n	93	33
Percentage of test data set (%)	73.81	26.19
Accuracy (%)	83.87	51.51
F_1-score	0.825	0.549

[a]PSE: patient safety event.

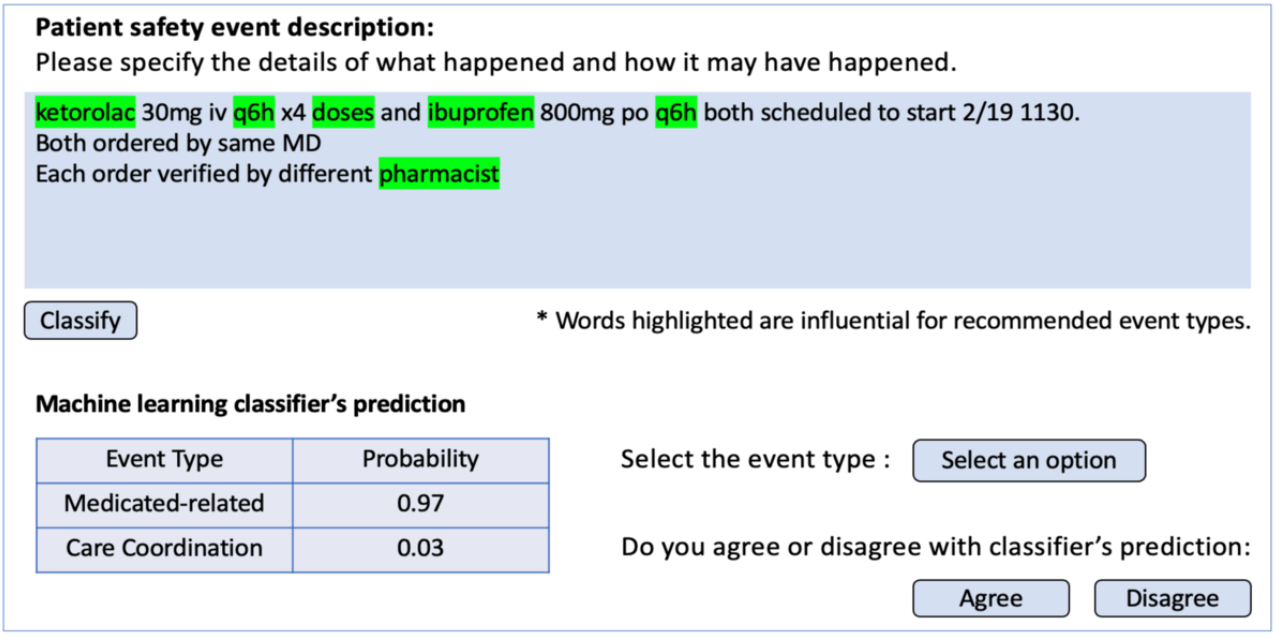
PSE Reporting System Interface

We designed an event reporting interface that integrates both the ML classifier and the LIME explainability technique. Figure 3 shows the event classification screen, where reporters enter a narrative description of the event after providing the details of the event, including date, time, unit, and information about the patient and reporter. Before describing the event in narrative form, reporters also choose among factors that contributed to the incident and the level of harm experienced by the patient. Once the reporter enters their narrative and selects the "classify" button, the system activates the ML classifier. Subsequently,

XSL•FO
RenderX

the interface displays the top 2 most probable event types, along with their associated probability distributions, in the lower left section. Simultaneously, the LIME technique will identify influential words that significantly contributed to the predicted event type, highlighting these words in green in the upper section of the dashboard. Based on the predicted event types and words highlighted for their influence on the prediction, the reporter may select the most suitable event type from a drop-down menu located in the lower-right section of the dashboard. Following this selection, reporters are queried on whether they agree with the classifier's prediction, and the collected data can be used to guide subsequent refinement of the ML classifier.

Figure 3. Interface visualization of a patient safety event report classifier coupled with the local interpretable model-agnostic explanations technique. MD: medical doctor.



Discussion

Overview

PSE event reporting systems are commonly used in health systems and hospitals across the world [46]. Data collected in PSE reporting systems drive quality improvement and patient safety efforts and supports regulatory reporting requirements for hospitals. The erroneous classification of PSE reports can impede the learning capabilities of the PSE reporting system, leading to suboptimal performance in detecting and preventing potential patient safety hazards [20]. It can also result in a substantial time cost for reclassifying PSE reports and compromise the integrity of a PSE database when analysts are investigating trends in events to develop effective solutions [17]. Previous studies have trained ML classifiers with static text representations for automatic PSE classification [12,17,25,26]. This study aimed to investigate whether using contextual text representations can further improve the accuracy of classifying PSE reports. We trained and evaluated a range of ML classifiers using both static and contextual text representations. To the best of our knowledge, this is the first time that contextual text representation has been used for training ML classifiers for PSE report classification. We analyzed the confusion matrix of the top-performing classifier to identify prevalent misclassified event types. Furthermore, aiming for more accurate and reliable PSE report classification, we incorporated an explainability technique to support human-AI collaboration and designed an interface to illustrate the possible integration of the ML classifier in PSE reporting systems.

Principal Findings

In this study, we extensively investigated the potential of using contextual representation for improving PSE report classification. The leading classifier trained with the static text representation (MLR trained with TF-IDF) was able to achieve an accuracy of 66.7% (84/126). This accuracy considerably exceeds the baseline accuracy of 29.4% (37/126), which involves classifying all PSE reports into the majority event type. However, using contextual text representation proved more efficacious. The SVM trained with contextual text representation (RoBERTa-base) was able to achieve an accuracy of 75.4% (95/126), reflecting a 13% (11/84) relative improvement in accuracy compared to the best-performing classifier trained with static text representation. While the achieved accuracy of 75.4% may not appear outstanding in isolation, it represents a significant advance compared with static text representations and exceeds the baseline, given the limited size of the data set. The improvement in classifier performance can be attributed to the use of contextual text representations, which can capture not only the meaning of individual words but also the complex and subtle ways in which words interact with each other in a specific context. Therefore, contextual text representation overcomes some limitations of static text representation, which relies primarily on word frequency and co-occurrence to represent text. Moreover, contextual text representation does not require explicit text normalization while also avoiding issues associated with high-dimensionality and sparsity commonly found in static

XSL•FO
RenderX

text representations. Hence, when training ML classifiers for PSE reporting systems, contextual text representation should be prioritized over static text representation to ensure the highest level of accuracy in classifying PSE reports.

As part of our investigation, we evaluated the performance of the top-performing classifier trained with contextual text representation on individual event types. While the classifier demonstrated impressive performance in accurately classifying laboratory test PSE reports (F_1-score=0.958), it struggled with classifying omissions or errors in assessment, diagnosis, and monitoring PSE reports, resulting in an unsatisfactory F_1-score of 0.400. To investigate this discrepancy, we analyzed the confusion matrix for the classifier and discovered that omissions or errors in assessment, diagnosis, and monitoring PSE reports were frequently misclassified as the coordination event type. This misclassification can be attributed to the multiclass nature of PSE reports. For example, a failure to document the removal of a patient's epidural catheter (omission or errors in assessment, diagnosis, and monitoring) could lead to a medication ordered by a physician (such as Lovenox) being withheld by the pharmacy due to a complication risk (coordination). On the other hand, the laboratory test is a more distinct event type in comparison to the other event types, and the classifier was able to correctly classify the majority of these reports. The observation obtained from the confusion matrix implies that PSE reports can potentially have more than 1 event type. This finding is consistent with previous studies [25,26]. The finding also underscores the need for further refinement in the development of the PSE taxonomy to create more distinctive event types. Another potential solution for addressing the multiclass nature of PSEs is to enable multiple event-type assignments [47]. Alternatively, the ML classifier can provide several probable event types, allowing the user to select the most appropriate one. We evaluated the top 2 accuracy of the top-performing ML classifier trained with contextual text representation and observed that 39% (12/31) of misclassified PSE reports' true event type was predicted as the second-highest probability event type by the classifier. The finding suggests that there is a greater chance for the ML classifier to provide the correct event type when considering multiple options. As event reporting systems usually encompass over 20 event types, which can be difficult to memorize or access [17], narrowing down the PSE report's potential event types to a smaller range also reduces the cognitive workload for PSE reporters during the classification process [48] and enhances the efficiency of reclassifying PSE reports for patient safety analysts.

We used LIME to showcase 3 predictions' explanations and demonstrated cases where the ML classifier used informative words for classifying the PSE report and where it used irrelevant words for classification. These results highlight the importance of not solely relying on the ML classifier's prediction and underscore the need for explainability and transparency in using the ML classifier for PSE report classification. Additionally, we showed the top 5 most prevalent words the ML classifier deemed important in the PSE reports for each event type. These words are indicative of the prevalent themes and issues within specific event types. Understanding the context and relationships between these prevalent informative words and specific event types can potentially provide valuable insights into the factors contributing to different types of PSEs. Furthermore, we have evaluated the top-performing ML classifier's performance on 2 subsets of PSE reports, differentiated by whether the highlighted word by LIME is relevant to the predicted event type. Our findings reveal that the majority of PSE reports (93/126) have at least 1 relevant word highlighted, with the classifier achieving an accuracy of 84% (78/93) on these reports. Conversely, accuracy drops to 52% (17/33) when irrelevant words are highlighted. Such a disparity in performance emphasizes the necessity for additional scrutiny from reporters and patient safety analysts, particularly when dealing with PSE reports that have irrelevant words highlighted.

While previous research has focused on the development of ML classifiers, none of these previous works have investigated the potential integration of the classifier within the PSE reporting system in a manner that aligns with the workflow of the front-end reporter. We designed an interface to demonstrate the feasibility of a collaborative human-AI approach for event categorization. The interface provides the PSE reporter with multiple probable event types and associated explanations for the ML classifier's prediction. This approach aligns with the principles of level 2 automation, where ML classifiers aids human decision-making rather than fully automating it [49]. This collaboration optimally combines human expertise with ML capabilities, potentially reducing cognitive workload and memorization of the taxonomy while also reducing the risks associated with overreliance on automation. Numerous studies have shown that the human-AI collaboration approach can improve the decision-making process [50-52], indicating its potential for enhancing PSE report classification. Furthermore, the interface also integrates the LIME explainability technique, which offers real-time insights into the rationale for the probable event types. Given the role of reporters and patient safety analysts in the incident reporting process, the use of explainability techniques can also increase trust in the recommendation provided by the ML classifier as it provides transparent and interpretable reasoning for the classification decisions [50,51]. Using LIME to highlight top informative words in real time for a PSE report can assist PSE reporters by emphasizing keywords in their narratives that are linked to the proposed classification. Highlighting informative words can also facilitate patient safety analysts working at the back end by providing insights into why a specific event type was chosen for classification. Such transparency not only clarifies current recommendations but also guides analysts in identifying influential terms for future report classifications. Previous research has illustrated the value of automation transparency in supporting appropriate levels of trust in the system, including decision support systems [32]. Additionally, regularly checking the explanations of the ML classifier's prediction enables continuous monitoring of the classifier's performance, identification of issues, and refinement [52]. As we have only designed the interface, additional research is needed to test the effectiveness of this approach in PSE report classification. Assessing the interface's impact on cognitive workload and decision-making accuracy is essential for ensuring its usability and adoption in the event reporting system. We plan to undertake

XSL•FO
RenderX

a usability testing study with health care professionals in a subsequent study.

Comparison With Previous Work

Research into the use of ML classifiers for the automation of PSE report classification has been relatively scarce. Wang et al [26] used logistic regression and SVM with the binary count, term frequency, and TF-IDF text representation to classify ten types of PSE reports, reaching an F_1-score as high as 0.783. However, they used a considerably larger data set (n=2860). Fong et al [17] achieved an accuracy rate of 92.0% (284/309) when they examined the usage of an ML classifier for classifying miscellaneous PSE reports using SVM, RF, and logistic regression with TF-IDF [17]. They also used a much larger data set (n=70,051). Ong et al [12] investigated the feasibility of using an ML classifier to automatically classify 2 types of PSE reports, including inadequate clinical handover and incorrect patient identification. They used Bag of Words model for text representation and trained both SVM and naive Bayes on classifying PSE reports, reaching accuracy as high as 98% (364/372). However, they framed the problem as a binary classification problem, which inherently has a higher baseline accuracy compared to our investigation. In this study, we've performed an in-depth comparative analysis with the available PSE data set and compared the established methods of classifying PSE reports and our novel method of using contextual text representations for classification. Our findings reveal that our proposed method outperforms the traditional models in terms of accuracy (ie, 84/126, 66.7% vs 95/126, 75.4%) and F_1-score (ie, 0.631 vs 0.753). This underlines the significance of our approach and its potential to advance the field of using ML classifiers for PSE report classification.

Limitations

There are several limitations to this study. First, the PSE reports used to train the ML classifiers were obtained from the maternal care units of a single hospital in the United States; therefore, the classifier might not generalize well to other settings. Second, this research's scope was constrained by the limited amount of PSE report data, and only 7 prevalent classes were incorporated for training the ML classifiers. The restricted quantity of PSE reports might also result in an underestimation of the ML classifier's actual capabilities [12]. Third, the quality of the LIME explanations was assessed by 2 graduate students; thus, further investigation is needed for a more robust validation of explanation quality. Furthermore, we have not yet empirically tested the interface for potential decision-making biases it may introduce.

Future research should investigate the performance of ML classifiers trained with contextual text representations on a larger and more diverse data set. Additionally, while we plan to refine the interface and test whether it supports event classification, future research can continue to investigate the appropriate way of incorporating the ML classifier into the reporting and reviewing workflow of PSE report classification and examine various human-AI collaboration approaches. Future studies should explore the potential biases (ie, automation bias) that the interface may introduce into the analysts' decision-making process.

Conclusions

Improving the precision of PSE report classifications is a multifaceted task, involving both the refinement of the event type taxonomy and adequate training of hospital staff on the event reporting system. Despite these challenges, ML classifiers offer substantial potential to support accurate classification throughout the reporting and reviewing process. The findings of this study contribute to the advancement of ML classifiers for PSE report classification by demonstrating the superior performance of contextual text representation over static text representations in achieving more accurate classification outcomes. The integration of explainability techniques in ML classifiers fosters trust in their usage and provides valuable insights for informed decision-making and potential adjustments to the classifier. An event reporting interface that integrates an ML classifier with collaborative decision-making capabilities offers the potential to achieve an efficient and reliable PSE report classification process. These approaches can ultimately help hospitals identify risks and hazards promptly and take timely and informed actions to mitigate adverse events and reduce patient harm.

Acknowledgments

This research was funded by the Agency for Healthcare Research and Quality (1R03HS027680).

Authors' Contributions

HC was responsible for the conceptualization, data analysis, and drafting of the manuscript. EC contributed to the conceptualization, methodology design, and review and revision of the document. DW contributed to data acquisition and funding acquisition. MA contributed to data acquisition, conceptualization, funding acquisition, and the review and revision of the manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Data preprocessing procedures.
[DOCX File , 14 KB - humanfactors_v11i1e53378_app1.docx]

XSL•FO
RenderX

Multimedia Appendix 2
Evaluation metrics for examining patient safety event machine learning classifiers.
[DOCX File , 15 KB - humanfactors_v11i1e53378_app2.docx]

Multimedia Appendix 3
The performance of machine learning classifiers in classifying patient safety event report event type.
[DOCX File , 20 KB - humanfactors_v11i1e53378_app3.docx]

References

1. Kohn LT, Corrigan JM, Donaldson MS, Committee on Quality of Health Care in America. To Err Is Human: Building a Safer Health System. Volume 627. Washington (DC): National Academies Press (US); 2000.
2. Nebeker JR, Barach P, Samore MH. Clarifying adverse drug events: a clinician's guide to terminology, documentation, and reporting. Ann Intern Med 2004 May 18;140(10):795-801 [FREE Full text] [doi: 10.7326/0003-4819-140-10-200405180-00009] [Medline: 15148066]
3. Chopard D, Treder MS, Corcoran P, Ahmed N, Johnson C, Busse M, et al. Text mining of adverse events in clinical trials: deep learning approach. JMIR Med Inform 2021;9(12):e28632 [FREE Full text] [doi: 10.2196/28632] [Medline: 34951601]
4. Makary MA, Daniel M. Medical error-the third leading cause of death in the US. BMJ 2016;353:i2139 [FREE Full text] [doi: 10.1136/bmj.i2139] [Medline: 27143499]
5. Slawomirski L, Auraaen A, Klazinga NS. The Economics of Patient Safety: Strengthening a Value-based Approach to Reducing Patient Harm at National Level. Paris: OECD Publishing; 2017.
6. Medication without harm. World Health Organization. 2023. URL: https://www.who.int/initiatives/medication-without-harm [accessed 2023-03-08]
7. Hewitt TA, Chreim S. Fix and forget or fix and report: a qualitative study of tensions at the front line of incident reporting. BMJ Qual Saf 2015;24(5):303-310 [FREE Full text] [doi: 10.1136/bmjqs-2014-003279] [Medline: 25749025]
8. Ward JK, Armitage G. Can patients report patient safety incidents in a hospital setting? a systematic review. BMJ Qual Saf 2012;21(8):685-699 [FREE Full text] [doi: 10.1136/bmjqs-2011-000213] [Medline: 22562875]
9. Li E, Clarke J, Ashrafian H, Darzi A, Neves AL. The impact of electronic health record interoperability on safety and quality of care in high-income countries: systematic review. J Med Internet Res 2022;24(9):e38144 [FREE Full text] [doi: 10.2196/38144] [Medline: 36107486]
10. Vincent C. Incident reporting and patient safety. BMJ 2007;334(7584):51 [FREE Full text] [doi: 10.1136/bmj.39071.441609.80] [Medline: 17218667]
11. Albolino S, Tartaglia R, Bellandi T, Amicosante AMV, Bianchini E, Biggeri A. Patient safety and incident reporting: survey of Italian healthcare workers. Qual Saf Health Care 2010;19(Suppl 3):i8-i12 [FREE Full text] [doi: 10.1136/qshc.2009.036376] [Medline: 20959324]
12. Ong MS, Magrabi F, Coiera E. Automated categorisation of clinical incident reports using statistical text classification. Qual Saf Health Care 2010;19(6):e55 [FREE Full text] [doi: 10.1136/qshc.2009.036657] [Medline: 20724392]
13. Herzer KR, Mirrer M, Xie Y, Steppan J, Li M, Jung C, et al. Patient safety reporting systems: sustained quality improvement using a multidisciplinary team and "good catch" awards. Jt Comm J Qual Patient Saf 2012;38(8):339-347 [FREE Full text] [doi: 10.1016/s1553-7250(12)38044-6] [Medline: 22946251]
14. Fong A, Hettinger AZ, Ratwani RM. Exploring methods for identifying related patient safety events using structured and unstructured data. J Biomed Inform 2015;58:89-95 [FREE Full text] [doi: 10.1016/j.jbi.2015.09.011] [Medline: 26432354]
15. Brubacher JR, Hunte GS, Hamilton L, Taylor A. Barriers to and incentives for safety event reporting in emergency departments. Healthc Q 2011;14(3):57-65 [FREE Full text] [doi: 10.12927/hcq.2011.22491] [Medline: 21841378]
16. Conceptual framework for the international classification for patient safety version 1.1: final technical report January 2009. World Health Organization. 2010. URL: https://apps.who.int/iris/handle/10665/70882 [accessed 2023-03-08]
17. Fong A, Behzad S, Pruitt Z, Ratwani RM. A machine learning approach to reclassifying miscellaneous patient safety event reports. J Patient Saf 2021;17(8):e829-e833 [FREE Full text] [doi: 10.1097/PTS.0000000000000731] [Medline: 32555052]
18. Ngo J, Lau D, Ploquin J, Receveur T, Stassen K, Del Castilho C. Improving incident reporting among physicians at south health campus hospital. BMJ Open Qual 2022;11(4):e001945 [FREE Full text] [doi: 10.1136/bmjoq-2022-001945] [Medline: 36207052]
19. Schrager JD, Schuler K, Isakov AP, Wright DW, Yaffee AQ, Jacobson KL, et al. Development and usability testing of a web-based COVID-19 self-triage platform. West J Emerg Med 2020;21(5):1054-1058 [FREE Full text] [doi: 10.5811/westjem.2020.7.48217] [Medline: 32970554]
20. Gong Y, Song HY, Wu X, Hua L. Identifying barriers and benefits of patient safety event reporting toward user-centered design. Saf Health 2015;1(1):1-9 [FREE Full text] [doi: 10.1186/2056-5917-1-7]
21. Gong Y. Data consistency in a voluntary medical incident reporting system. J Med Syst 2011;35(4):609-615 [FREE Full text] [doi: 10.1007/s10916-009-9398-y] [Medline: 20703528]

XSL•FO
RenderX

22. Koike D, Ito M, Horiguchi A, Yatsuya H, Ota A. Implementation strategies for the patient safety reporting system using consolidated framework for implementation research: a retrospective mixed-method analysis. BMC Health Serv Res 2022;22(1):409 [FREE Full text] [doi: 10.1186/s12913-022-07822-9] [Medline: 35346182]
23. Biannual incident report. Clinical Excellence Commission. URL: https://www.cec.health.nsw.gov.au/Review-incidents/Biannual-Incident-Report [accessed 2023-05-16]
24. Organisation patient safety incident report up to March 2022. NHS England. 2022. URL: https://www.england.nhs.uk/publication/organisation-patient-safety-incident-report-up-to-march-2022/ [accessed 2023-03-08]
25. Evans HP, Anastasiou A, Edwards A, Hibbert P, Makeham M, Luz S, et al. Automated classification of primary care patient safety incident report content and severity using supervised Machine Learning (ML) approaches. Health Informatics J 2020;26(4):3123-3139 [FREE Full text] [doi: 10.1177/1460458219833102] [Medline: 30843455]
26. Wang Y, Coiera E, Runciman W, Magrabi F. Using multiclass classification to automate the identification of patient safety incident reports by type and severity. BMC Med Inform Decis Mak 2017;17(1):84 [FREE Full text] [doi: 10.1186/s12911-017-0483-8] [Medline: 28606174]
27. Minaee S, Kalchbrenner N, Cambria E, Nikzad N, Chenaghlu M, Gao J. Deep learning—based text classification: a comprehensive review. ACM Comput Surv 2021;54(3):1-40 [FREE Full text] [doi: 10.1145/3439726]
28. Liu Q, Kusner MJ, Blunsom P. A survey on contextual embeddings. ArXiv Preprint posted online on April 13, 2020. [doi: 10.48550/arXiv.2003.07278]
29. Asan O, Bayrak AE, Choudhury A. Artificial intelligence and human trust in healthcare: focus on clinicians. J Med Internet Res 2020;22(6):e15154 [FREE Full text] [doi: 10.2196/15154] [Medline: 32558657]
30. Sharma M, Savage C, Nair M, Larsson I, Svedberg P, Nygren JM. Artificial intelligence applications in health care practice: scoping review. J Med Internet Res 2022;24(10):e40238 [FREE Full text] [doi: 10.2196/40238] [Medline: 36197712]
31. Sezgin E. Artificial intelligence in healthcare: complementing, not replacing, doctors and healthcare providers. Digit Health 2023;9:20552076231186520 [FREE Full text] [doi: 10.1177/20552076231186520] [Medline: 37426593]
32. Hemmer P, Schemmer M, Riefle L, Rosellen N, Vössing M, Kühl N. Factors that influence the adoption of human-AI collaboration in clinical decision-making. ArXiv Preprint posted online on April 19, 2022. [doi: 10.48550/arXiv.2204.09082]
33. Amann J, Blasimme A, Vayena E, Frey D, Madai VI, Precise4Q consortium. Explainability for artificial intelligence in healthcare: a multidisciplinary perspective. BMC Med Inform Decis Mak 2020;20(1):310 [FREE Full text] [doi: 10.1186/s12911-020-01332-6] [Medline: 33256715]
34. Devlin J, Chang MW, Lee K, Toutanova K. BERT: pre-training of deep bidirectional transformers for language understanding. Minneapolis, Minnesota: Association for Computational Linguistics; 2019 Presented at: 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies; June 2-7, 2019; Minneapolis, Minnesota p. 4171-4186.
35. Pennington J, Socher R, Manning C. GloVe: global vectors for word representation. Doha, Qatar: Association for Computational Linguistics; 2014 Presented at: 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP); October 25-29, 2014; Doha, Qatar p. 1532-1543. [doi: 10.3115/v1/d14-1162]
36. Models. Hugging Face. 2022. URL: https://huggingface.co/models [accessed 2023-03-09]
37. Liu Y, Ott M, Goyal N, Du J, Joshi M, Chen D, et al. RoBERTa: a robustly optimized BERT pretraining approach. ArXiv Preprint posted online on July 26 2019. [doi: 10.5260/chara.21.2.8]
38. Kumar P, Bhatnagar R, Gaur K, Bhatnagar A. Classification of imbalanced data:review of methods and applications. IOP Conf Ser Mater Sci Eng 2021;1099:012077 [FREE Full text] [doi: 10.1088/1757-899x/1099/1/012077]
39. Chawla NV, Bowyer KW, Hall LO, Kegelmeyer WP. SMOTE: synthetic minority over-sampling technique. J Artif Int Res 2002;16(1):321-357 [FREE Full text] [doi: 10.1613/jair.953]
40. Wang L, Zhang Y, Chignell M, Shan B, Sheehan KA, Razak F, et al. Boosting delirium identification accuracy with sentiment-based natural language processing: mixed methods study. JMIR Med Inform 2022;10(12):e38161 [FREE Full text] [doi: 10.2196/38161] [Medline: 36538363]
41. Ribeiro MT, Singh S, Guestrin C. "Why should I trust you?": explaining the predictions of any classifier. 2016 Presented at: KDD '16: The 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining; August 13-17, 2016; San Francisco, California p. 1135-1144 URL: https://dl.acm.org/doi/abs/10.1145/2939672.2939778 [doi: 10.1145/2939672.2939778]
42. Palojoki S, Saranto K, Reponen E, Skants N, Vakkuri A, Vuokko R. Classification of electronic health record-related patient safety incidents: development and validation study. JMIR Med Inform 2021;9(8):e30470 [FREE Full text] [doi: 10.2196/30470] [Medline: 34245558]
43. The collaborative interface design tool. Figma. URL: https://www.figma.com/ [accessed 2023-09-24]
44. Delio J, Catalanotti JS, Marko K, Paul C, Taffel M, Ho G, et al. Integrating adverse event reporting into a free-text mobile application used in daily workflow increases adverse event reporting by physicians. Am J Med Qual 2020;35(5):374-379 [FREE Full text] [doi: 10.1177/1062860619891995] [Medline: 31795736]
45. Klemp K, Zwart D, Hansen J, Hellebek T, Luettel D, Verstappen W, et al. A safety incident reporting system for primary care. A systematic literature review and consensus procedure by the LINNEAUS collaboration on patient safety in primary care. Eur J Gen Pract 2015;21(sup1):39-44 [FREE Full text] [doi: 10.3109/13814788.2015.1043728] [Medline: 26339835]

XSL•FO
RenderX

46. Hasegawa T, Fujita S. Patient safety policies experiences, effects and priorities; lessons from OECD member states (version 2.0). Patient Safety Global Ministerial Summit 2018. 2018. URL: https://www.mhlw.go.jp/psgms2018/pdf/document/5_Document.pdf [accessed 2023-12-21]
47. Wang Y, Coiera E, Runciman W, Magrabi F. Automating the identification of patient safety incident reports using multi-label classification. Stud Health Technol Inform 2017;245:609-613. [Medline: 29295168]
48. Buettner R. Cognitive workload of humans using artificial intelligence systems: towards objective measurement applying eye-tracking technology. In: KI 2013: Advances in Artificial Intelligence Berlin. Heidelberg: Springer Berlin Heidelberg; 2013 Presented at: 36th Annual German Conference on AI, Koblenz; September 16-20, 2013; Germany p. 37-48. [doi: 10.1007/978-3-642-40942-4_4]
49. Berretta S, Tausch A, Ontrup G, Gilles B, Peifer C, Kluge A. Defining human-AI teaming the human-centered way: a scoping review and network analysis. Front Artif Intell 2023;6:1250725 [FREE Full text] [doi: 10.3389/frai.2023.1250725] [Medline: 37841234]
50. Lai Y, Kankanhalli A, Ong DC. Human-AI collaboration in healthcare: a review and research agenda. 2021 Presented at: 54th Hawaii International Conference on System Sciences (HICSS); January 5-8, 2021; Grand Wailea, Maui, Hawaii URL: https://scholarspace.manoa.hawaii.edu/server/api/core/bitstreams/495a50c0-16e9-4c3e-a00b-34aeec9dff6b/content [doi: 10.24251/hicss.2021.046]
51. Reverberi C, Rigon T, Solari A, Hassan C, Cherubini P, GI Genius CADx Study Group, et al. Experimental evidence of effective human-AI collaboration in medical decision-making. Sci Rep 2022;12(1):14952 [FREE Full text] [doi: 10.1038/s41598-022-18751-2] [Medline: 36056152]
52. Lee MH, Siewiorek DP, Smailagic A, Bernardino A, Bermúdez i Badia S. A human-AI collaborative approach for clinical decision making on rehabilitation assessment. New York, NY, USA: Association for Computing Machinery; 2021 Presented at: CHI '21: CHI Conference on Human Factors in Computing Systems; May 8-13, 2021; Yokohama Japan p. 1-14 URL: https://dl.acm.org/doi/10.1145/3411764.3445472 [doi: 10.1145/3411764.3445472]

Abbreviations

AI: artificial intelligence
KNN: k-nearest neighbor
LIME: local interpretable model-agnostic explanations
ML: machine learning
MLR: multinomial logistic regression
OECD: Organization for Economic Cooperation and Development
PSE: patient safety event
RF: random forest
SVM: support vector machine
TF-IDF: term frequency–inverse document frequency

Edited by E Borycki, A Kushniruk; submitted 06.10.23; peer-reviewed by H Tibble, S Matathil; comments to author 19.11.23; revised version received 30.11.23; accepted 03.12.23; published 25.01.24.

Please cite as:
Chen H, Cohen E, Wilson D, Alfred M
A Machine Learning Approach with Human-AI Collaboration for Automated Classification of Patient Safety Event Reports: Algorithm Development and Validation Study
JMIR Hum Factors 2024;11:e53378
URL: https://humanfactors.jmir.org/2024/1/e53378
doi:10.2196/53378
PMID:38271086

©Hongbo Chen, Eldan Cohen, Dulaney Wilson, Myrtede Alfred. Originally published in JMIR Human Factors (https://humanfactors.jmir.org), 25.01.2024. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on https://humanfactors.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Testing the Feasibility and Acceptability of Using an Artificial Intelligence Chatbot to Promote HIV Testing and Pre-Exposure Prophylaxis in Malaysia: Mixed Methods Study

Min Hui Cheah[1], MClinPharm; Yan Nee Gan[2], MSc; Frederick L Altice[3,4,5,6], MA, MD; Jeffrey A Wickersham[3,5,6], PhD; Roman Shrestha[5,6,7], MPH, PhD; Nur Afiqah Mohd Salleh[2], PhD; Kee Seong Ng[1], MD, PhD; Iskandar Azwa[1,6], MD; Vimala Balakrishnan[8], PhD; Adeeba Kamarulzaman[9], MBBS; Zhao Ni[5,10], PhD, BMed

[1]Department of Medicine, Faculty of Medicine, University of Malaya, Kuala Lumpur, Malaysia
[2]Department of Social and Preventive Medicine, Faculty of Medicine, University of Malaya, Kuala Lumpur, Malaysia
[3]Section of Infectious Disease, Department of Internal Medicine, Yale School of Medicine, New Haven, CT, United States
[4]Division of Epidemiology of Microbial Diseases, Yale School of Public Health, New Haven, CT, United States
[5]Center for Interdisciplinary Research on AIDS (CIRA), Yale University, New Haven, CT, United States
[6]Centre of Excellence for Research in AIDS (CERiA), Faculty of Medicine, University of Malaya, Kuala Lumpur, Malaysia
[7]Department of Allied Health Sciences, University of Connecticut, Storrs, CT, United States
[8]Department of Information Systems, Faculty of Computer Science and Information Technology, University of Malaya, Kuala Lumpur, Malaysia
[9]Monash University Malaysia, Kuala Lumpur, Malaysia
[10]School of Nursing, Yale University, Orange, CT, United States

Corresponding Author:
Zhao Ni, PhD, BMed
School of Nursing
Yale University
400 West Campus Drive
Orange, CT, 06477
United States
Phone: 1 203 737 3039
Email: zhao.ni@yale.edu

Abstract

Background: The HIV epidemic continues to grow fastest among men who have sex with men (MSM) in Malaysia in the presence of stigma and discrimination. Engaging MSM on the internet using chatbots supported through artificial intelligence (AI) can potentially help HIV prevention efforts. We previously identified the benefits, limitations, and preferred features of HIV prevention AI chatbots and developed an AI chatbot prototype that is now tested for feasibility and acceptability.

Objective: This study aims to test the feasibility and acceptability of an AI chatbot in promoting the uptake of HIV testing and pre-exposure prophylaxis (PrEP) in MSM.

Methods: We conducted beta testing with 14 MSM from February to April 2022 using Zoom (Zoom Video Communications, Inc). Beta testing involved 3 steps: a 45-minute human-chatbot interaction using the think-aloud method, a 35-minute semistructured interview, and a 10-minute web-based survey. The first 2 steps were recorded, transcribed verbatim, and analyzed using the Unified Theory of Acceptance and Use of Technology. Emerging themes from the qualitative data were mapped on the 4 domains of the Unified Theory of Acceptance and Use of Technology: performance expectancy, effort expectancy, facilitating conditions, and social influence.

Results: Most participants (13/14, 93%) perceived the chatbot to be useful because it provided comprehensive information on HIV testing and PrEP (*performance expectancy*). All participants indicated that the chatbot was easy to use because of its simple, straightforward design and quick, friendly responses (*effort expectancy*). Moreover, 93% (13/14) of the participants rated the overall chatbot quality as high, and all participants perceived the chatbot as a helpful tool and would refer it to others. Approximately 79% (11/14) of the participants agreed they would continue using the chatbot. They suggested adding a local language (ie, Bahasa Malaysia) to customize the chatbot to the Malaysian context (*facilitating condition*) and suggested that the chatbot should also incorporate more information on mental health, HIV risk assessment, and consequences of HIV. In terms of *social influence*, all

XSL•FO
RenderX

participants perceived the chatbot as helpful in avoiding stigma-inducing interactions and thus could increase the frequency of HIV testing and PrEP uptake among MSM.

Conclusions: The current AI chatbot is feasible and acceptable to promote the uptake of HIV testing and PrEP. To ensure the successful implementation and dissemination of AI chatbots in Malaysia, they should be customized to communicate in Bahasa Malaysia and upgraded to provide other HIV-related information to improve usability, such as mental health support, risk assessment for sexually transmitted infections, AIDS treatment, and the consequences of contracting HIV.

(JMIR Hum Factors 2024;11:e52055) doi:10.2196/52055

KEYWORDS

artificial intelligence; acceptability; chatbot; feasibility; HIV prevention; HIV testing; men who have sex with men; MSM; mobile health; mHealth; preexposure prophylaxis; PrEP; mobile phone

Introduction

Background

HIV continues to be a global health concern causing approximately 630,000 deaths yearly worldwide [1]. In 2019, approximately 62% of the new HIV infections among adults worldwide occurred within key populations and their sexual partners [2]. Men who have sex with men (MSM) accounted for 23% of new infections of HIV, which was much higher than the percentage of new infections in other key populations, such as people who used drugs (10%), sex workers (8%), and transgender people (2%) in 2019 [3]. Malaysia is a Southeast Asian country with a population of 33.5 million, with 1 in 5 MSM living with HIV [4]. Over the past 2 decades, the mode of HIV transmission in Malaysia has shifted from needle sharing to sexual transmission, particularly among MSM [5].

HIV testing is a prerequisite to effective HIV prevention and early treatment initiation [6]; people at risk for HIV or seropositive individuals need to be tested for HIV before being linked to health care services [7]. Despite the importance of HIV testing, it is disproportionately lower among MSM in Malaysia [8]. New HIV testing guidelines recommend that MSM at high risk for HIV should be tested every 3 to 6 months, but most MSM in Malaysia do not test optimally. Studies in Malaysia have found that only 9.5% of MSM tested more than once a year. In Malaysia, engaging in same-sex sexual behavior is prohibited by both secular and Sharia laws, leading to significant levels of stigma and discrimination within society [9]. As a result, many MSM may be hesitant or unwilling to engage with health care providers and outreach workers. Therefore, designing new strategies to promote HIV testing among MSM in Malaysia is urgently needed [10].

Using portable electronic devices with software programs to deliver health care services and manage patient information is known as mobile health (mHealth) [11]. mHealth interventions could reduce barriers to HIV testing for MSM by reducing in-person contact and offering internet-based platforms for HIV testing [12,13]. Studies have demonstrated that mHealth interventions using smartphones and apps could increase the uptake of HIV testing while protecting the privacy of MSM [14-16], and MSM in Malaysia have a high acceptance of the use of mHealth for HIV testing and prevention [13,17,18]. Recent breakthroughs in artificial intelligence (AI) and machine learning can potentially automate and scale up these mHealth interventions through chatbots, a computer program that can mimic human conversation [19]. However, leveraging chatbot technology to promote HIV testing and prevention is in its infancy [15,20]. Although chatbot technology holds immense potential to prevent HIV, a lack of research in this field undermines its significance. The creation of ChatGPT has brought attention to the significance of studying chatbot technology for health care.

Our team has conducted formative research to understand HIV prevention chatbots in Malaysia and has identified the perceived benefits, limitations, and preferred features of AI chatbots for HIV testing and prevention among MSM [13]. On the basis of the study findings, we developed an HIV prevention AI chatbot prototype named Haris (a common Malaysian name) and a website called MYHIV365 (*MY* symbolizes Malaysia, *HIV* implies health care services aimed at preventing HIV, and *365* indicates the services are available every day of the year). Haris was hosted on MYHIV365 and could provide information on the 3 themes most needed by MSM: HIV testing, mental health, and pre-exposure prophylaxis (PrEP). PrEP is a highly effective HIV prevention method that involves the use of antiretroviral medication by at-risk individuals to prevent getting HIV from sex or injection drug use. Haris imitates human intelligence and can interact with users to provide support, including ordering free HIV self-testing kits, screening for depression, and recommending MSM-friendly clinics where individuals can get tested for HIV and receive PrEP.

Objectives

Despite the meticulous design and alpha testing (internal testing) of Haris among professors, experts, and community advisory board members, its feasibility and acceptability in preventing HIV among MSM is still unknown. Therefore, we conducted beta testing (testing in a real-world environment by actual users) of Haris among 14 MSM in Malaysia to address this knowledge gap. Specifically, we examined the use of the AI chatbot for delivering health information and improving linkage to HIV testing, PrEP, and care. We also investigated key strategies to refine the feasibility and acceptability of the AI chatbot in this study.

XSL•FO
RenderX

Methods

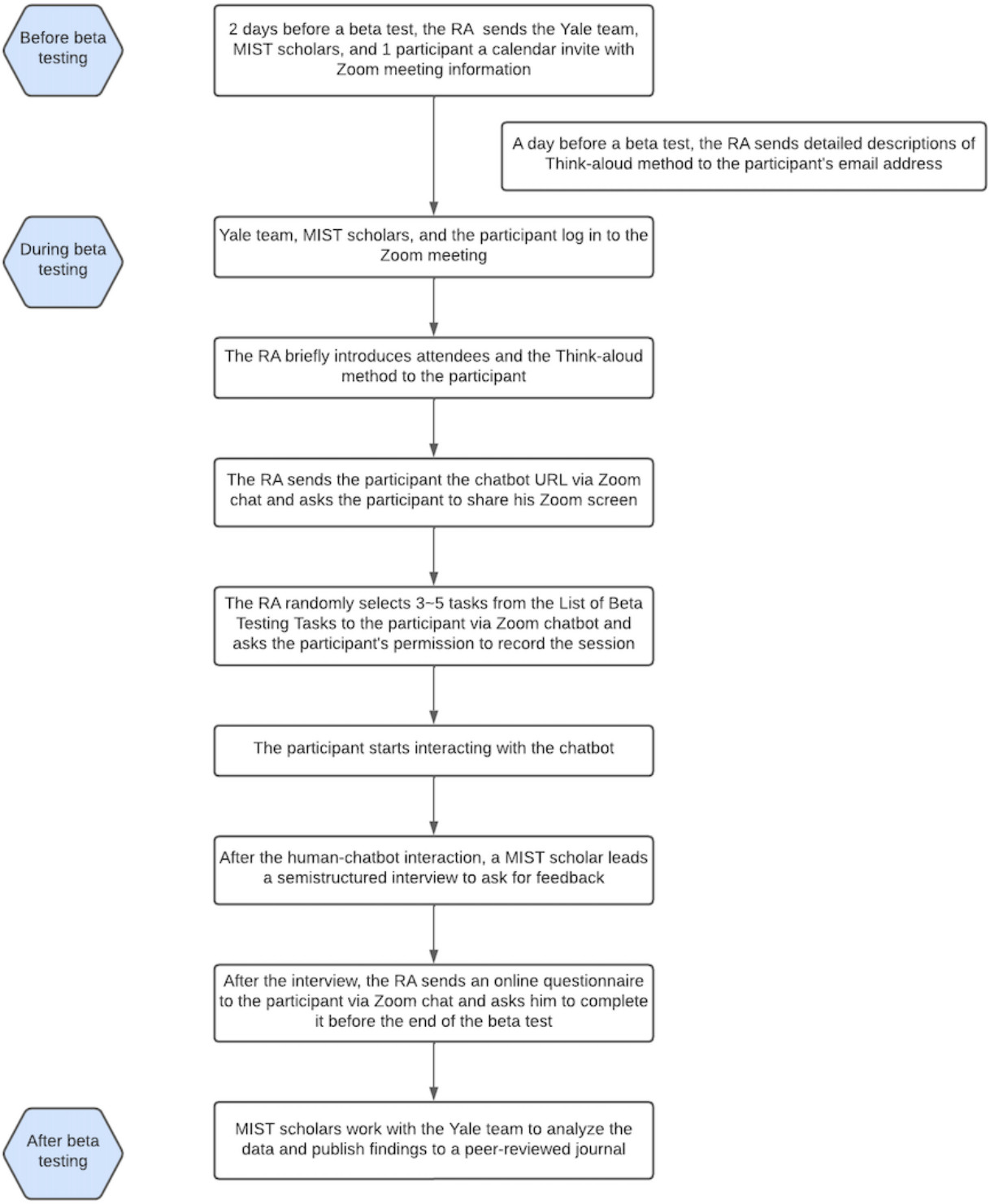
Study Design and Participants

Beta testing of the AI chatbot prototype was conducted with 14 MSM by an experienced qualitative interviewer (ZN) with expertise in chatbot development and HIV prevention in Malaysia and 4 trainees in the Malaysian Implementation Science Training program (Fogarty International Center, D43TW011324). Participants were recruited in Malaysia from February to April 2022 via social networking apps commonly used by MSM, including Grindr, Hornet, Blued, and WhatsApp. The procedures for participant recruitment have been published elsewhere [13]. A web-based screener including questions on demographic characteristics and HIV prevention practices was used. The eligibility criteria included (1) self-identification as a cisgender man, (2) age ≥18 years, (3) condomless sex with another man in the past 6 months, and (4) being HIV negative or of unknown status.

Each beta test involved the following three steps: (1) a 45-minute human-chatbot interaction using the think-aloud method [21]; (2) a 35-minute semistructured interview; and (3) a 10-minute web-based survey. The first 2 steps were conducted via Zoom (Zoom Video Communications, Inc), recorded, and transcribed verbatim. Specifically, 2 days before the test, a research assistant sent a calendar invite with Zoom meeting information to the interviewer and participant. One day before beta testing, the research assistant emailed the participant a detailed description of the human-chatbot interaction (Multimedia Appendix 1). During the human-chatbot interaction, participants were asked to share their screen via Zoom and access the chatbot through a URL sent by the research assistant. After the participants obtained access to the chatbot, the research assistant randomly selected 3 to 5 tasks from the list of beta testing tasks (Multimedia Appendix 2) and asked the participants to complete them through the chatbot. Some examples of the tasks include “find a clinic that can provide HIV testing service in Kuala Lumpur” and “find out the common symptoms of depression through the chatbot.” The study procedure is described in Figure 1.

XSL•FO
RenderX

Figure 1. Study procedure. MIST: Malaysian Implementation Science Training; RA: research assistant.



After the human-chatbot interaction, we conducted a semistructured interview (Multimedia Appendix 3 [22]) soliciting participants' feedback on two themes: (1) experience navigating the chatbot and (2) how the chatbot should be made available to a wider audience. During the interview, participants were asked several questions regarding their experience with the chatbot, such as "How was your experience with the AI chatbot?", "What feature of the AI chatbot do you like the most?", and "What information needs to be added to the AI chatbot to increase its popularity among MSM?" After the interviews concluded, the participants were provided with a survey link to assess the feasibility and acceptability of the AI chatbot. The feasibility of the chatbot was measured through 4 outcomes, including participants' ratings of the chatbot's quality, satisfaction, intention to continue using the chatbot, and willingness to refer it to others. The outcomes were measured using a 10-point rating scale, with higher scores indicating more favorable outcomes (Multimedia Appendix 4). For example,

XSL•FO
RenderX

participants' satisfaction with the chatbot was measured by using the question, "How satisfied were you with the experience of interacting with the chatbot?" The score of "0" stands for not satisfactory at all and "10" stands for extremely satisfactory. The acceptability of the chatbot was measured using the standardized System Usability Scale [23] and an adjusted Chatbot Usability Scale [24]. The combination of the 2 scales provided a comprehensive evaluation of the acceptability of our chatbot.

Ethical Considerations

The participants provided electronic consent before initiating the beta testing. This study was approved by the institutional review board of Yale University (approval #2000027864) and Medical Research Ethics Committee of the University of Malaya (approval #2021112-10729). This research was conducted in accordance with the ethical standards of the 1964 Helsinki Declaration and its later amendments or comparable ethical standards.

Conceptual Framework for Analysis

The Unified Theory of Acceptance and Use of Technology (UTAUT) was used as a conceptual framework to guide the analysis of the experience of MSM using the AI chatbot for HIV testing and prevention in Malaysia. UTAUT consists of four domains: (1) performance expectancy, (2) effort expectancy, (3) facilitating conditions, and (4) social influence [25]. The definitions of these 4 domains have been published elsewhere [13]. UTAUT was chosen for the following reasons. First, this AI chatbot was developed based on the findings from a formative research project that was analyzed using UTAUT [13]. Therefore, using the same theory, we can compare the results of the 2 studies on the 4 domains and are more likely to find out the feasibility and acceptability of the AI chatbot. Second, UTAUT emphasizes user-centered perspective, which allows researchers to assess the acceptance of the AI chatbot from the users' perception. Third, UTAUT has been extensively used to identify users' acceptance of technology and was reported to be effective and of high validity [26,27].

Analyses

All transcripts were cross-checked for accuracy and completeness by 7 researchers (MHC, YNG, NAMS, KSN, ZN, and 2 research assistants). Each of the 7 researchers independently coded 2 transcripts using NVivo 10 software (QSR International), compiled codes, and mapped the emerging themes from the qualitative data on the 4 domains of UTAUT, including performance expectancy, effort expectancy, facilitating conditions, and social influence. Discrepancies in codes and themes were addressed in group discussions where there was discordance in coding. We ceased the qualitative analysis when the results reached saturation, and no new themes emerged. The participants' quotes are presented throughout the results with additional quotes given in Multimedia Appendix 5. Quantitative data from the survey were analyzed using SAS (version 9.4; SAS Institute) and are presented as descriptive statistics.

Results

Participant Characteristics

The 14 participants were on average in their mid-20s (mean 25.6, SD 4.2 years), and most of them (13/14, 93%) used smartphones as the primary means to access the internet. Most participants (10/14, 71%) were Malay, followed by Chinese (3/14, 21%) and Indian (1/14, 7%). About one-third of the participants (5/14, 36%) had taken PrEP previously, and only 14% (2/14) of them were currently taking PrEP. The demographic characteristics are summarized in Table 1.

XSL•FO
RenderX

Table 1. Participant demographic details (N=14).

Characteristics	Values
Age (y), mean (SD)	25.6 (4.2)
Ethnicity, n (%)	
Malay	10 (71)
Chinese	3 (21)
Indian	1 (7)
Sexual orientation, n (%)	
Bisexual	3 (21)
Gay	11 (79)
Employment status, n (%)	
Student	6 (43)
Working full time	8 (57)
Highest level of education, n (%)	
Diploma or bachelor degree	8 (57)
Master degree or PhD	3 (21)
Secondary school	3 (21)
Average monthly income (MYR[a]; 1 MYR=US $0.21), n (%)	
<2000	6 (43)
2000-4000	5 (36)
>4000	3 (21)
Daily access to the internet, n (%)	
Yes	14 (100)
Primary device for accessing the internet, n (%)	
Smartphone	13 (93)
Laptop computer	1 (7)
Had ever taken PrEP[b], n (%)	
Yes	5 (36)
No	9 (64)
Currently taking PrEP, n (%)	
Yes	2 (14)
No	12 (86)

[a]MYR: Malaysian Ringgit.

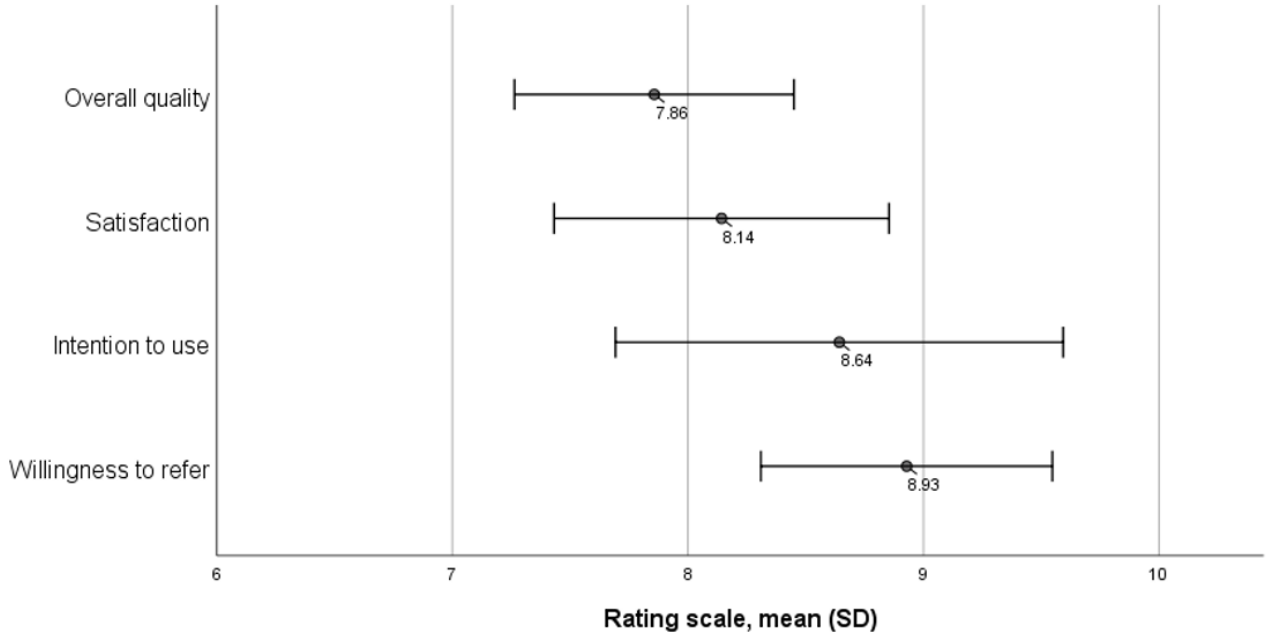
[b]PrEP: pre-exposure prophylaxis.

Feasibility

The mean scores on the 4 metrics of the feasibility of the chatbot, overall quality, satisfaction, intention to continue using, and willingness to refer to others were 7.86 (SD 1.03), 8.14 (SD 1.23), 8.64 (SD 1.65), and 8.93 (SD 1.07), respectively, on a scale from 0 to 10 (Figure 2).

XSL•FO
RenderX

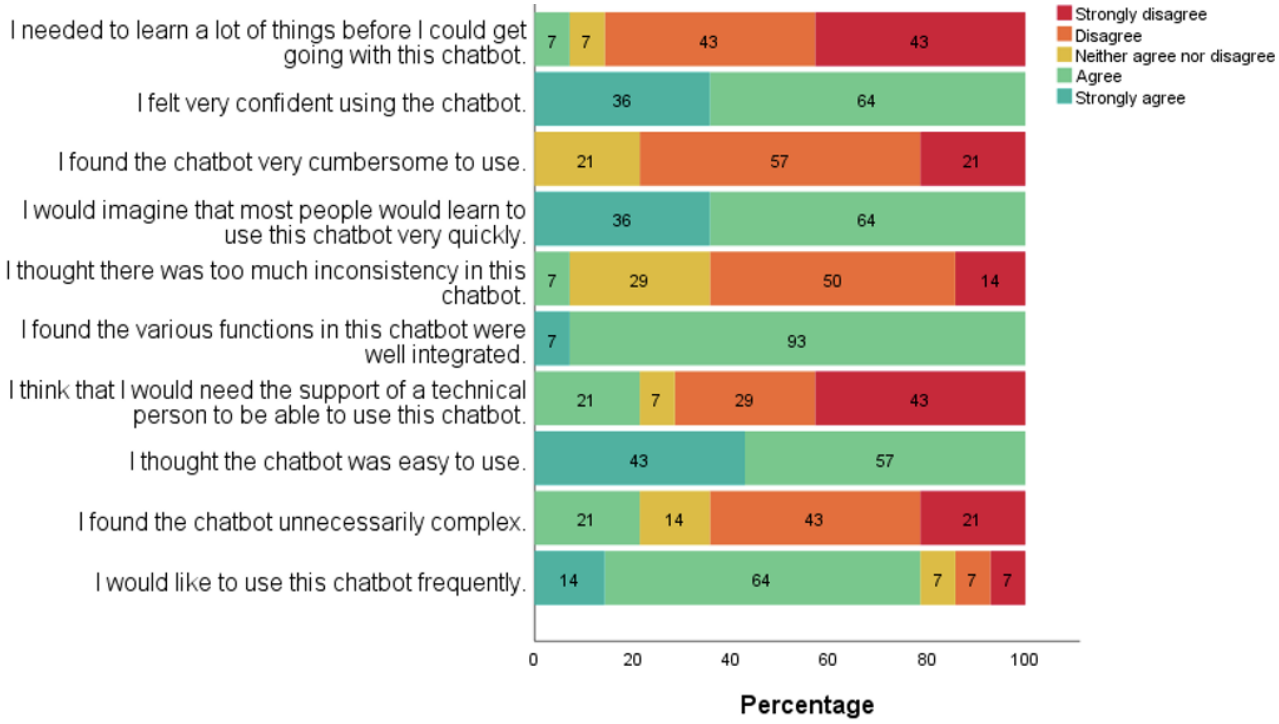
Figure 2. Feasibility ratings of the chatbot.



Acceptability

The participants found the chatbot acceptable, as it was perceived as easy to navigate and capable of providing valuable information (Multimedia Appendix 6). Specifically, all participants (14/14, 100%) expressed confidence in using the chatbot, believing that others could also quickly master its use (Figure 3). The overall mean (SD) score of the System Usability Scale was 76.07 (SD 8.19), which is greater than the recommended acceptable cutoff score of 68 [23].

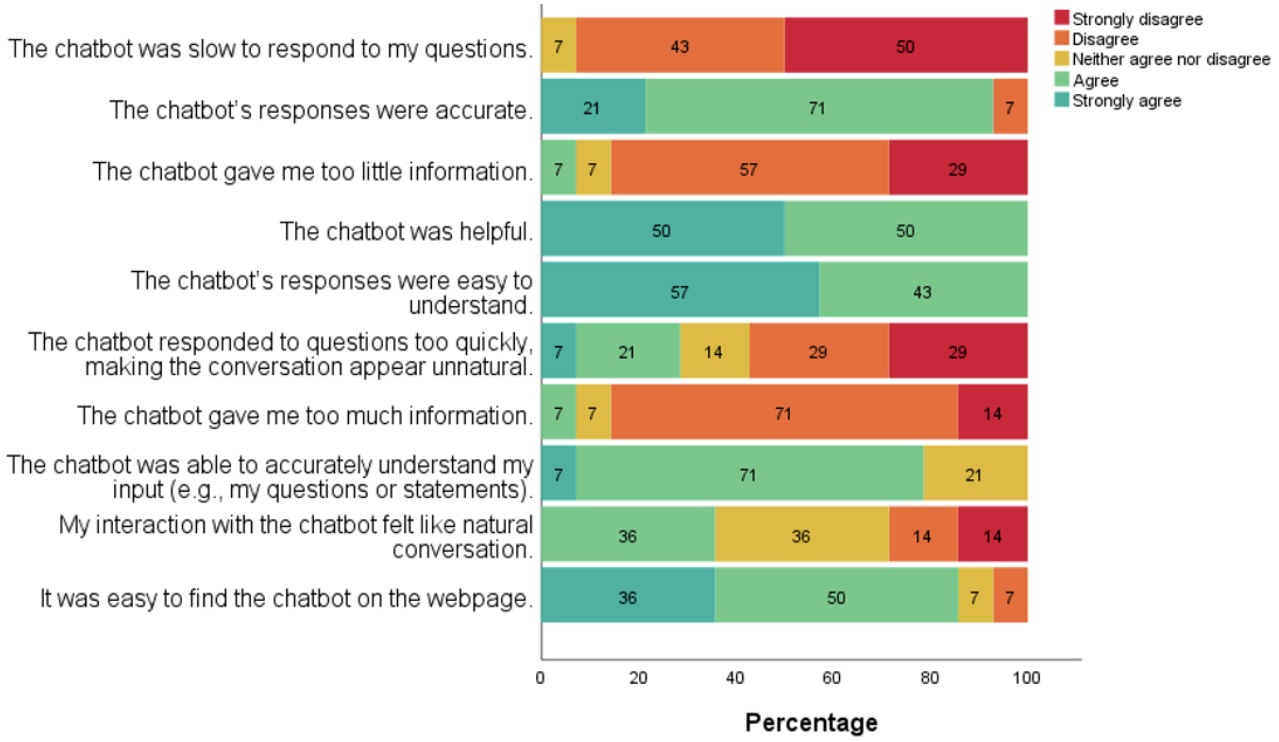
Figure 3. System Usability Scale outcomes.



Moreover, 78% (11/14) of the participants agreed that the chatbot could understand their inputs accurately (Figure 4). However, only 36% (5/14) of the participants agreed that their interaction with the chatbot felt like a natural conversation.

XSL•FO
RenderX

Figure 4. Adjusted Chatbot Usability Scale outcomes.



In addition to measuring feasibility and acceptability, we summarized the study findings based on the 4 domains of UTAUT: performance expectancy, effort expectancy, facilitating conditions, and social influence.

Performance Expectancy

Overall Perception

Participants responded positively about the performance of the AI chatbot. Quantitative data analysis revealed that all participants (14/14, 100%) perceived the chatbot as a helpful tool, they would refer it to others, and 93% (13/14) of them highly rated the overall quality of the chatbot. These results were consistent with the qualitative finding that participants were satisfied with the information provided by the chatbot and deemed it a trustable source.

Contributors to Positive Performance Expectancy

The chatbot served as a reliable source of information. Participants expressed their trust in this chatbot as they believed it was developed to help people learn more about HIV. For instance, one participant stated that he trusted the chatbot more than other internet-based platforms as the following:

> *I can actually trust, trust this chatbot more than I can trust the Internet.* [Interview E]

Most participants expressed that the information provided by the chatbot was comprehensive and satisfactory. The same participant highlighted the following:

> *...everything is there, everything is informative...* [Interview E]

Participants described the chatbot's function of ordering free HIV self-test kits as one of the most useful functions. The simple and straightforward instructions from the chatbot significantly encouraged participants to perform HIV self-testing and helped prevent misuse of the self-test kit. A participant elaborated on the following:

> *...it never crossed my mind that you can do HIV self-test just using it (test stick) over the gums and without blood...* [Interview G]

In addition to the positive feedback on the chatbot's ability to order free HIV self-test kits, participants also expressed appreciation for the information provided by the chatbot on MSM-friendly local clinics where they could test for HIV and receive PrEP consultation. For example, the participant stated the following:

> *That was beyond amazing...they give [me] the addresses, contact numbers. So, I would say, if a person really needs to do the [HIV] testing, essential information like that would be very useful, so I think it's more than helpful. The options [of HIV testing clinics] are not just [limited to] one or two [clinics], you know, so that's good.* [Interview G]

The other participants stressed that the chatbot's features related to HIV self-testing and venue-based HIV testing were complementary and appreciated having access to both options through the chatbot. Although performing self-testing at home may be convenient, it may lack the human interaction that some MSM need for support during the testing process. By offering both self-testing and venue-based testing options, the chatbot gave MSM the flexibility to choose the option that best suited their individual needs and preferences. For example, a participant emphasized that the chatbot could enable his MSM friends to conduct HIV self-testing and then see a health care professional for advice on complex and sensitivity issues:

XSL•FO
RenderX

> *...the reason why [they] went to a health care facility was [that they] can have someone tell them that you know, "Being tested positive for HIV [was] not the end of the world. To reduce the [HIV] symptoms, to reduce the [HIV] effect, to reduce [their HIV] viral load... [they] can still live a normal life and so on," which might be something that the self-test kits [were] lacking.* [Interview C]

Along with the positive feedback on the chatbot's features relevant to HIV testing, the participants also expressed their favorable feelings toward the PrEP information provided by the chatbot. For example, a participant mentioned that the chatbot could send him introductory information to allow him to comfortably assess his risk level and help him decide if he needed to take PrEP. Another participant reiterated the first participant's point by emphasizing the difficulties that MSM in determining if they should take PrEP, stating the following:

> *A lot of people [MSM] are always asking themselves, "Should I get PrEP? Am I at risk? Do I need to take PrEP as a precautionary measure?"... So these are [questions] that MSM usually a bit too scared to ask the doctors.* [Interview C]

Major Concerns

Participants suggested that the chatbot would be improved if it could provide more information and resources relevant to mental health, as mental health issues were the prominent problems that MSM faced in Malaysia. Participants wanted the chatbot to provide information on strategies for managing stress, statistics about depression among MSM, peer consultation for depression, and professional health care services to prevent and treat depression. The participants also highlighted that the MYHIV365 website, where the chatbot was embedded, should provide more resources related to mental health. For example, one participant described this problem as follows:

> *The website did not have links to any information regarding mental health issues, and that is a glaring issue for it to be left out like that.* [Interview A]

The same statement was echoed by another participant, who stressed the following:

> *I just find that for the mental health, it's kind of short.* [Interview G]

Relevant Features Suggested by Participants That Are Needed to Improve the Chatbot

Although the chatbot made it easy for participants to receive an HIV self-test kit, one participant suggested that a step-by-step video demonstrating how to use the kit could be helpful for MSM who were testing for HIV. The participant stated the following:

> *If the self-test kit has got like an instructional video or something like that to kind of guide the users along the way of getting [themselves] tested, I think that'd be great because not everyone knows how to use a kit successfully.* [Interview C]

Although the AI chatbot was developed primarily for HIV prevention and to assist with HIV testing and access to PrEP, the participants pointed out that some participants may test positive for HIV and would benefit from learning more about accessing HIV care and related antiretroviral therapy services. In addition, participants suggested that the chatbot should provide more information about antiretroviral therapy so that users could better manage HIV by knowing potential drug interactions and side effects. The participants also recommended providing more information about high-risk behaviors and sexually transmitted diseases to help increase awareness about HIV and sexually transmitted diseases among MSM. One participant stated the following:

> *I think [providing more information about] HIV treatment would [be] very helpful because those who might be exposed to HIV would definitely want to know what the treatment is all about.* [Interview E]

This participant's statement was echoed by another participant, who stated the following:

> *...HIV and STDs...[are] not the same, but...I thought [they were] the same...I thought HIV and STDs were not curable...so I think it will be great if you can add STDs [to the chatbot].* [Interview E]

Effort Expectancy

Overall Perception

All surveyed participants (14/14, 100%) agreed that the chatbot was easy to use, and 86% (12/14) of the participants were satisfied with the chatbot. In the qualitative interviews, participants reported consistent feedback that the chatbot was user friendly and convenient to use, and they were satisfied with the chatbot because of its simple, straightforward design and quick, friendly responses. However, they were concerned about the technical issues, including the address input and text alignments (refer to the *Major Concerns* section). The participants also felt that tailoring the chatbot to the local context and adding a "human touch" would be helpful.

Positive Contributors to Low Effort Expectancy

Many participants expressed satisfaction with the chatbot because of its prompt response, expert information, and plain interface. Two participants commented the following:

> *...white and blue colors [are] neutral, and it [the chatbot] takes into account [of] color blindness as well, so that's great.* [Interview C]
>
> *...[I] got a quick response [from the chatbot].* [Interview I]

The individualized and user-centered features of the chatbot, which cater to users with different levels of communication skills, were highlighted among the participants. For instance, one participant stated that the chatbot offered an ideal platform for MSM who are less comfortable interacting with others. A participant stated the following:

> *As we all know, some of us didn't have the skills to communicate, so I think...[the] chatbot… will definitely help. I think it was great.* [Interview E]

XSL•FO
RenderX

Moreover, participants thought the chatbot was useful as it facilitated them to obtain culturally tailored health information. The chatbot met users' needs by providing a menu of options for users to choose from. Compared with obtaining health information in clinical settings in Malaysia, the chatbot was much simpler. A participant elaborated on the following:

> *When [the chatbot] come[s] up with three options like that, I can explore myself...I would say that [the chatbot] is more precise; it gets to the point directly.* [Interview G]

The health intervention being tailored to the local setting was highly valued by the participants. Responses from the AI chatbot that contained localized features, specifically the use of "Manglish," a less formal form of Malaysian English, were appreciated by several participants. The feature of "Manglish," which was not in the standardized form of English, has added a local flavor to the AI chatbot, which some participants found amusing. A participant stated the following:

> *The impression that this chatbot...probably comes from America. It's in English, so the moment it puts up a Malaysian style saying "Boleh"... I'm very amused with this [style].* [Interview G]

Major Concerns

Some participants spoke about the difficulty in filling in their home addresses using the current prompts on the AI chatbot when they needed to order an HIV self-test kit; the chatbot required a step-by-step input of addresses, which was counterintuitive and inefficient. Participants preferred the standard address format in Malaysia over the step-by-step input format, in which incomplete addresses would triage further prompts to ask participants to refill the HIV self-test order. For example, one participant stated the following:

> *In Malaysia, we don't use the term "line address" or "street address". We usually enter the full address with the postcode and then the city and state. The one on the chatbot seems to be how addresses are filled in the United States. That part needs to be tweaked slightly based on Malaysian cities.* [Interview C]

In addition, participants expressed that the address of the clinics provided by the chatbot needed to be tailored to Malaysian culture. For example, the district options may only be needed for certain states in Malaysia. A participant stated the following:

> *I think depends on the size of the state...we don't have to call out (provide choices for) all the districts because Perlis is already small enough, and I think...people can just go easily from one place to another in Perlis. But if...it is a big state...we need to divide it using district.* [Interview I]

Relevant Features Suggested by Participants That Are Needed to Improve the Chatbot

Although participants were satisfied with the AI chatbot, 2 participants suggested that the chatbot's interface could be improved by adding more spaces between sentences, and the alignment of sentences should be adjusted to make the chatbot look more professional. Two participants described the following:

> *...everything is tightly together with very little gap...there should be proper spacing...* [Interview L]
>
> *The text is not properly centered in some of the boxes, and I feel like it could [be] a better design to make it look more professional.* [Interview A]

The use of English as the only language of the AI chatbot was perceived by participants as a barrier to implementing the chatbot in Malaysia. Although all participants were proficient in English, concerns arose for the communities where English was not widely spoken. Participants suggested that the chatbot should be able to communicate in Bahasa Malaysia or Mandarin, given that the 2 languages are widely spoken in Malaysia. A participant stated the following:

> *...perhaps to have another option of language...I think that would be able to cover more people within the local population.* [Interview C]

Adding a "human touch" to the chatbot can create a more engaging and user-friendly experience for the users interacting with the chatbot. The participant described the following:

> *...ideally, we would want [the chatbot's response] to be as human as possible, and not so robotic in its responses...a nice touch to make someone feel slightly comfortable.* [Interview C]

Facilitating Conditions

Overall Perception

Participants reported 2 major facilitating conditions for the use of the AI chatbot. First, the social distancing policy adopted by the Malaysian government during the COVID-19 pandemic significantly increased the use of internet-based platforms to seek health information and consult about health issues among MSM. The participants expressed that the AI chatbot was a novel tool to promote HIV testing and prevention among MSM in Malaysia. A participant highlighted the convenience of using the chatbot as an alternative to meeting health care workers during the COVID-19 pandemic as follows:

> *...because now it's COVID, everyone is doing it in IT (information technology) format. Having an AI chatbot is definitely much more convenient than meeting people...* [Interview G]

Second, the AI chatbot's capability of referring webpages to participants where they could find mental health information, community support, and counselors was a significant facilitator for them to accept the chatbot. Many participants stated that it was much easier to obtain information through the links provided by the chatbot than searching for information via websites or mobile apps. A participant stated the following:

> *When you interact with it (the chatbot), it throws out links to you. It's easier to navigate to the particular links from there.* [Interview G]

XSL•FO
RenderX

Relevant Features Suggested by Participants That Are Needed to Improve the Chatbot

Participants suggested that the chatbot could be promoted through social media platforms, such as Facebook, Twitter, Instagram, YouTube, TikTok, and Telegram because these platforms were widely used by MSM as sources of information. Among all social networking apps, participants stated that Twitter was the best platform to advertise the AI chatbot because Twitter enabled users to post clickable links in the comments section where other users could access the chatbot. Participants further reported that Telegram was a more suitable platform for hosting the chatbot than the most popular text messaging app in Malaysia, WhatsApp. Telegram offers a more private and secure environment for MSM to ask questions or express concerns about HIV and AIDS. Participants also suggested that building a trustable relationship between the AI chatbot and the MSM community is key to implementing the AI chatbot in Malaysia. Given that there were many scams through pop-up advertisements on social media platforms, a participant described the following imagined scenario:

> *...if we play our Facebook, Twitter, Instagram, or YouTube, there are always mini advertisements, so who knows, [whether we] can add this [AI chatbot]?...I need to know about this [chatbot], and I hope this [chatbot] is not a scam.* [Interview B]

Social Influence

In terms of *social influence*, the chatbot was perceived as helpful in avoiding HIV stigma and thus could increase the HIV testing rate and PrEP uptake frequency. Quantitative data analysis found that 79% (11/14) of the participants agreed to continue using the chatbot. During the interviews, these participants reported that societal stigma and discrimination related to HIV and AIDS would make them more likely to use the chatbot. They expressed discomfort in asking people questions about HIV and AIDS as they were afraid of encountering stigma and negative attitudes from others. MSM often preferred to seek information through internet-based platforms, and the chatbot was helpful, particularly for people living in small social circles. A participant elaborated on the following:

> *...this topic [HIV] is quite sensitive to most people, it will create like a negative energy around you...personally I don't go and ask people what HIV is, I will search myself maybe on the Internet...* [Interview E]

The societal stigma and discrimination toward HIV and AIDS also facilitated participants to select HIV self-testing at home rather than testing in a clinical setting. Many participants appreciated that the chatbot offered them an opportunity to receive free HIV self-test kits while protecting their privacy. Two participants who used to be shy about discussing HIV described the following:

> *Because from the MSM community, some of us are not very comfortable of getting [HIV] test kits on site, because like...fear of the stigma, that the society will judge.* [Interview D]

> *I can directly book the test kit through the chatbot, which is very useful and informative…my identity will remain anonymous, so people don't know me.* [Interview E]

Participants deemed the AI chatbot useful and expressed their willingness to recommend the AI chatbot to others. Some participants suggested that the chatbot should be promoted among MSM who frequently use social networking apps, such as Grindr, Hornet, and Blued, to find sexual partners because those MSM were at higher risk for HIV and had greater need for HIV information. A participant stated the following:

> *I have the impression that anyone would actually need it [HIV testing]. But if we look at it from another angle, people on hookup apps like Grindr have a high tendency to hook up using those apps compared to those who don't use them...we need to introduce the chatbot to them because...they...[have] been highly exposed [to HIV].* [Interview G]

Discussion

Principal Findings

The feasibility and acceptability of leveraging AI chatbots to promote HIV testing and PrEP among MSM in Malaysia is high. Discrimination and stigma toward HIV and AIDS are major barriers for MSM to access high-quality HIV testing and prevention services in Malaysia, and they are also primary facilitators for MSM to seek health information via internet-based platforms. Our AI chatbot prototype provides a platform for MSM to order free HIV self-test kits in an MSM-friendly environment and to empower them with resources and instructions. MSM who prefer to interact with health care providers in person can also locate HIV testing clinics or PrEP clinics through the AI chatbot. MSM highlighted these functions of the AI chatbot as very useful.

Similar to other studies, AI chatbots were well received by users [28,29]. An AI chatbot could enhance engagement with the key population [30]. As contemporary social patterns increasingly involve the integration of AI into everyday routines, AI chatbots could contribute to delivering precise details regarding HIV testing to individuals actively seeking such information. A chatbot named Eli, developed by the United Nations Educational, Scientific, and Cultural Organization, received highly favorable user feedback and was widely acclaimed [29]. Eli offers a range of services, including details on HIV prevention, testing, and treatment and assistance in overcoming fears and concerns. Compared with Eli, our AI chatbot did not have information on treatment for AIDS and provided limited mental health support. Integrating these functions into our AI chatbot may support its usability. Nevertheless, our AI chatbot offers free HIV self-test kits and locates local clinics in Malaysia for HIV testing, PrEP consultation, and mental health care.

From our previous formative research, we know that factors facilitating the acceptance of an HIV prevention AI chatbot include providing useful information and having the capacity to solve problems [13]. In this study, participants reported that our AI chatbot was able to provide useful information and help

XSL•FO
RenderX

solve problems. This was indicated by the results that all participants perceived this chatbot as a helpful tool, and most participants deemed the chatbot a reliable source of information with a high satisfaction score. However, one area that required significant improvement in the chatbot was its conversation flow, as only 36% (5/14) of the participants felt that their interaction with the chatbot resembled a natural conversation. This was similar to another study where the quick response of the chatbot was deemed not humanlike and perceived as a disadvantage [28]. To address this issue and advance the chatbot, improving its algorithm and continuing training it using AI and machine learning techniques based on feedback from a larger sample size is crucial. Considering that the use of AI chatbots in health care is still in its early stages, this finding holds particular significance for designing AI chatbots. To enhance usability and promote the implementation of AI chatbots in health care, the chatbots must possess the ability to initiate natural conversations with humanlike characteristics. In addition, they should be equipped to effectively address users' questions and concerns while ensuring the security and safety of users' information.

The chatbot's plain interface and simple design were popular among MSM. Digital health interventions are useful, but knowing how to navigate a digital system sometimes could be daunting for users. Through this study, we are clear that accurate and simple responses without errors and redundant information were key to the acceptability of AI chatbots among MSM. Our participants reported that the AI chatbot helped them avoid societal stigma and protected their privacy, which increased their acceptability of using the chatbot to test for HIV. This finding is consistent with our previous formative research finding that addressing sociocultural barriers can facilitate the acceptance of an AI chatbot [13]. The chatbot does not require users to provide registration information. Therefore, it can maintain participants' anonymity. However, it is still necessary for the chatbot to clarify to users that the backend researchers and engineers who have access to users' conversations and information will not expose users' information to others. This suggestion is consistent with our study findings and some previous studies showing that mHealth interventions could improve HIV testing rates if users' anonymity were guaranteed [14,17].

Some technical-related issues negatively affected the participants' experience of navigating the chatbot. The inconvenient address input process and repeated steps owing to incomplete information contributed to the inconsistency and complexity of the chatbot, prompting participants to seek technical assistance. Many of these resulted from cultural differences, as the address options were designed based on overseas settings. This signifies the importance of tailoring the chatbot to the local context to improve usability. In addition, using localized language could also enhance the participants' satisfaction with the chatbot. Despite the challenges inherent in adopting novel technology, the advantages of using chatbots to connect with high-risk populations could significantly impact the efforts to address public health emergencies.

In line with other studies conducted in Malaysia, MSM are keen to peruse the information on PrEP and mental health, particularly the information on where the PrEP and mental health clinics are located [14]. Most participants in our study felt that they would like to see more information through the chatbot introducing AIDS, its treatment, mental health issues, and sexually transmitted infections to better understand and manage AIDS, including how to prevent high-risk behaviors and where to seek timely help [12]. In Malaysia, professional and MSM-friendly care for mental health needs to be developed as most MSM reported that culturally sensitive information and resources regarding mental health issues were difficult to obtain. Interestingly, researchers have identified several obstacles to the adoption of AI chatbots for mental health care among users [31]. These include concerns related to privacy risks, restricted conversational engagement, negative user perceptions of personality traits (such as rudeness, lack of empathy, patronization, and being judgmental), and a lack of trust in the app's creators. Nevertheless, Eli chatbot overcame all these challenges by having a language that merges expertise and respect for the user, ensuring speech that is gender neutral and devoid of stigmatizing elements [29]. Our AI chatbot also had a similar language as Eli, which warrants future support on mental health issues.

To increase the use of the AI chatbot, it needs to be embedded in social media platforms that MSM frequently use. The geosocial networking apps where MSM find sexual partners, such as Grindr, Hornet, and Blued, and websites owned by nongovernmental organizations and MSM-friendly clinics are important venues to advertise the AI chatbot. MSM preferred these platforms because they are trusted and frequently used by MSM. Dissemination of the AI chatbot should be promoted among young MSM who use geosocial networking apps to find sexual partners because they are at a higher risk for HIV. Through this study, we found that to embed an AI chatbot into an internet-based platform for health promotion, researchers and engineers must consider the platform's characteristics, including its target population, level of privacy, and user-friendliness. Findings from this study will be used to improve the AI chatbot before testing on a larger scale through a national observational study in Malaysia. AI chatbots are a promising tool for promoting HIV testing and prevention. The AI chatbot must be made visible to MSM to increase its usability among MSM. Adopting the right dissemination strategies is key to increasing the visibility of AI chatbots and bringing significant impact to the MSM community. In addition, it is important for researchers to consider the sustainability of AI chatbots for MSM care in a context where sex-same sexual behaviors are criminalized. The policies and laws in Malaysia pose significant challenges on the sustainability of leveraging AI and securing funding for MSM care research. In such a political environment, it is crucial for researchers to collaborate with local nongovernmental organization and MSM-friendly clinics that operate within the existing Malaysian legal framework. Future research should focus on developing innovative and culturally tailored AI interventions to combat HIV among MSM, promote public health in Malaysia, and advocate for changes in discriminatory policies and laws to enhance the testing, implementation, and sustainability of these AI interventions.

XSL•FO
RenderX

Limitations

Testing the AI chatbot among its end users (ie, MSM) was an important step in determining its feasibility and acceptability in Malaysia and collecting feedback to improve the chatbot further. Although this study contributed important scientific knowledge, it had several limitations. One of the limitations is that we only included MSM who can read English, as the AI chatbot is currently only available in this language. Thus, the reach of the AI chatbot may be limited only to those fluent in English, which is not the case for most MSM in Malaysia. Therefore, our findings may not be generalizable to MSM who cannot read English. Considering that Malaysia is a multilingual country with Bahasa Malaysia as the official language, the chatbot must be improved to communicate in Bahasa Malaysia or Mandarin to reach a wider audience and promote greater access to HIV self-testing and PrEP. In addition, our participants were highly educated; this may lead to bias as they might possess a certain level of knowledge and health literacy, thus facilitating their interactions with the chatbot. Therefore, the findings may differ in the less educated or literate group. In addition, our study only included MSM aged ≥18 years; therefore, the study findings do not capture the perceptions of younger MSM who are typically more tech-savvy and susceptible to HIV. Although obtaining consent from younger MSM in Malaysia for HIV-related research is a significant challenge, future studies should consider conducting surveys and interviews with MSM aged <18 years who can provide insights into the experiences and needs of the younger MSM.

Conclusions

The AI chatbot was found to be feasible and acceptable among MSM, highlighting features, such as being informative, being able to respond to users' questions, and having a simple and user-friendly interface. Adapting the AI chatbot to local cultures, including support for other languages, and providing additional information such as mental health support, risk assessment for sexually transmitted infections, AIDS treatment, and the consequences of contracting HIV would contribute to the successful implementation and dissemination of the AI chatbot in Malaysia.

Acknowledgments

This work was supported by research grants from the National Institutes of Health (R21TW011663 and R33TW011663 for ZN and FLA; R21AI152927 for JAW, FLA, and ZN; and D43TW011324 for FLA) and the Fund for Lesbian and Gay Studies Research Award from Yale University for ZN.

Authors' Contributions

MHC and ZN wrote the first draft of the manuscript. YNG, NAMS, and KSN assisted in both developing the qualitative analysis and writing the qualitative results. All authors contributed to the development of the study protocol, revised the subsequent version of the manuscript, and approved the submitted version.

Conflicts of Interest

None declared.

Multimedia Appendix 1
An email that the research assistant sent to participants introducing the human-chatbot interaction.
[PDF File (Adobe PDF File), 64 KB - humanfactors_v11i1e52055_app1.pdf]

Multimedia Appendix 2
The list of beta testing tasks.
[PDF File (Adobe PDF File), 478 KB - humanfactors_v11i1e52055_app2.pdf]

Multimedia Appendix 3
The guide on chatbot beta testing.
[PDF File (Adobe PDF File), 336 KB - humanfactors_v11i1e52055_app3.pdf]

Multimedia Appendix 4
The scales used for measuring outcome variables.
[PDF File (Adobe PDF File), 195 KB - humanfactors_v11i1e52055_app4.pdf]

Multimedia Appendix 5
Participants' insights with illustrative quotes.
[DOCX File , 32 KB - humanfactors_v11i1e52055_app5.docx]

Multimedia Appendix 6

XSL•FO
RenderX

Screenshots demonstrating the chatbot's simple conversation flow.
[PDF File (Adobe PDF File), 958 KB - humanfactors_v11i1e52055_app6.pdf]

References

1. HIV and AIDS. World Health Organization. 2022. URL: https://www.who.int/news-room/fact-sheets/detail/hiv-aids [accessed 2023-12-26]
2. New HIV infections increasingly among key populations. Joint United Nations Programme on HIV/Aids. 2020. URL: https://www.unaids.org/en/resources/presscentre/featurestories/2020/september/20200928_new-hiv-infections-increasingly-among-key-populations [accessed 2023-12-26]
3. Seizing the moment: tackling entrenched inequalities to end epidemics. Joint United Nations Programme on HIV/Aids. 2020. URL: http://tinyurl.com/4zyedprp
4. Suleiman A, Tze CP. Country progress report on HIV/AIDS 2019 Malaysia. HIV/STI/Hepatitis, C Section of Ministry of Health, Malaysia. 2019. URL: https://www.moh.gov.my/moh/resources/Penerbitan/Laporan/Umum/Report_GAM_2019_(Final).pdf [accessed 2023-12-22]
5. Global AIDS monitoring: country progress report-Malaysia. Ministry of Health, Malaysia. 2021. URL: https://www.moh.gov.my/moh/resources/Penerbitan/Laporan/Umum/MYS_country_report_2021.pdf [accessed 2023-12-26]
6. Montaner JS, Lima VD, Harrigan PR, Lourenço L, Yip B, Nosyk B, et al. Expansion of HAART coverage is associated with sustained decreases in HIV/AIDS morbidity, mortality and HIV transmission: the "HIV treatment as prevention" experience in a Canadian setting. PLoS One 2014;9(2):e87872 [FREE Full text] [doi: 10.1371/journal.pone.0087872] [Medline: 24533061]
7. Gardner EM, McLees MP, Steiner JF, Del Rio C, Burman WJ. The spectrum of engagement in HIV care and its relevance to test-and-treat strategies for prevention of HIV infection. Clin Infect Dis 2011 Mar 15;52(6):793-800 [FREE Full text] [doi: 10.1093/cid/ciq243] [Medline: 21367734]
8. Risher K, Mayer KH, Beyrer C. HIV treatment cascade in MSM, people who inject drugs, and sex workers. Curr Opin HIV AIDS 2015 Nov;10(6):420-429 [FREE Full text] [doi: 10.1097/COH.0000000000000200] [Medline: 26352393]
9. Shrestha R, Wickersham JA, Khati A, Azwa I, Ni Z, Kamarulzaman A, et al. Clinic-integrated mobile health intervention ("JomPrEP" App) to improve uptake of HIV testing and pre-exposure prophylaxis among men who have sex with men in Malaysia: protocol for an intervention development and multiphase trial. JMIR Res Protoc 2022 Dec 21;11(12):e43318 [FREE Full text] [doi: 10.2196/43318] [Medline: 36542425]
10. Beyrer C, Sullivan PS, Sanchez J, Dowdy D, Altman D, Trapence G, et al. A call to action for comprehensive HIV services for men who have sex with men. Lancet 2012 Jul 28;380(9839):424-438 [FREE Full text] [doi: 10.1016/S0140-6736(12)61022-8] [Medline: 22819663]
11. Ni Z, Wu B, Yang Q, Yan LL, Liu C, Shaw RJ. An mHealth intervention to improve medication adherence and health outcomes among patients with coronary heart disease: randomized controlled trial. J Med Internet Res 2022 Mar 09;24(3):e27202 [FREE Full text] [doi: 10.2196/27202] [Medline: 35262490]
12. Aliabadi N, Carballo-Dieguez A, Bakken S, Rojas M, Brown 3rd W, Carry M, et al. Using the information-motivation-behavioral skills model to guide the development of an HIV prevention smartphone application for high-risk MSM. AIDS Educ Prev 2015 Dec;27(6):522-537 [FREE Full text] [doi: 10.1521/aeap.2015.27.6.522] [Medline: 26595265]
13. Peng ML, Wickersham JA, Altice FL, Shrestha R, Azwa I, Zhou X, et al. Formative evaluation of the acceptance of HIV prevention artificial intelligence chatbots by men who have sex with men in Malaysia: focus group study. JMIR Form Res 2022 Oct 06;6(10):e42055 [FREE Full text] [doi: 10.2196/42055] [Medline: 36201390]
14. Shrestha R, Lim SH, Altice FL, Copenhaver M, Wickersham JA, Saifi R, et al. Use of smartphone to seek sexual health information online among Malaysian men who have sex with men (MSM): implications for mHealth intervention to increase HIV testing and reduce HIV risks. J Community Health 2020 Feb;45(1):10-19 [FREE Full text] [doi: 10.1007/s10900-019-00713-x] [Medline: 31375976]
15. Marcus JL, Sewell WC, Balzer LB, Krakower DS. Artificial intelligence and machine learning for HIV prevention: emerging approaches to ending the epidemic. Curr HIV/AIDS Rep 2020 Jun;17(3):171-179 [FREE Full text] [doi: 10.1007/s11904-020-00490-6] [Medline: 32347446]
16. Shrestha N, Kukkonen-Harjula KT, Verbeek JH, Ijaz S, Hermans V, Pedisic Z. Workplace interventions for reducing sitting at work. Cochrane Database Syst Rev 2018 Jun 20;6(6):CD010912 [FREE Full text] [doi: 10.1002/14651858.CD010912.pub4] [Medline: 29926475]
17. Shrestha R, Maviglia F, Altice FL, DiDomizio E, Khati A, Mistler C, et al. Mobile health technology use and the acceptability of an mHealth platform for HIV prevention among men who have sex with men in Malaysia: cross-sectional respondent-driven sampling survey. J Med Internet Res 2022 Jul 25;24(7):e36917 [FREE Full text] [doi: 10.2196/36917] [Medline: 35877172]
18. Rosen AO, Wickersham JA, Altice FL, Khati A, Luces JR, Ni Z, et al. Adaptation of a theory-based, clinic-affiliated smartphone app to improve HIV testing and pre-exposure prophylaxis uptake among gay, bisexual, and other men who have sex with men in Malaysia. iproc 2022 Jul 6;8(1):e39358 [FREE Full text] [doi: 10.2196/39358]

XSL•FO
RenderX

19. Caldarini G, Jaf S, McGarry K. A literature survey of recent advances in chatbots. Information 2022 Jan 15;13(1):41 [FREE Full text] [doi: 10.3390/info13010041]
20. Ahlström MG, Ronit A, Omland LH, Vedel S, Obel N. Algorithmic prediction of HIV status using nation-wide electronic registry data. EClinicalMedicine 2019 Dec;17:100203 [FREE Full text] [doi: 10.1016/j.eclinm.2019.10.016] [Medline: 31891137]
21. Eccles DW, Arsal G. The think aloud method: what is it and how do I use it? Qual Res Sport Exerc Health 2017 May 26;9(4):514-531 [FREE Full text] [doi: 10.1080/2159676x.2017.1331501]
22. Nick B. Everything you need to know about beta testing. Adobe. 2019. URL: https://xd.adobe.com/ideas/process/user-testing/everything-you-need-to-know-about-betatesting/ [accessed 2023-12-26]
23. System usability scale (SUS). U.S. General Services Administration. URL: https://www.usability.gov/how-to-and-tools/methods/system-usability-scale.html [accessed 2023-09-22]
24. Borsci S, Malizia A, Schmettow M, van der Velde F, Tariverdiyeva G, Balaji D, et al. The chatbot usability scale: the design and pilot of a usability scale for interaction with AI-based conversational agents. Pers Ubiquit Comput 2021 Jul 21;26(1):95-119 [FREE Full text] [doi: 10.1007/s00779-021-01582-9]
25. Venkatesh V, Morris MG, Davis GB, Davis FD. User acceptance of information technology: toward a unified view. MIS Q 2003;27(3):425-478 [FREE Full text] [doi: 10.2307/30036540]
26. Venkatesh V, Thong JY, Xu X. Consumer acceptance and use of information technology: extending the unified theory of acceptance and use of technology. MIS Q 2012;36(1):157-178 [FREE Full text] [doi: 10.2307/41410412]
27. Gupta B, Dasgupta S, Gupta A. Adoption of ICT in a government organization in a developing country: an empirical study. J Strateg Inf Syst 2008 Jun;17(2):140-154 [FREE Full text] [doi: 10.1016/j.jsis.2007.12.004]
28. Ntinga X, Musiello F, Keter AK, Barnabas R, van Heerden A. The feasibility and acceptability of an mHealth conversational agent designed to support HIV self-testing in south Africa: cross-sectional study. J Med Internet Res 2022 Dec 12;24(12):e39816 [FREE Full text] [doi: 10.2196/39816] [Medline: 36508248]
29. Chatbot answers young people's questions about HIV, health and relationships. Joint United Nations Programme on HIV/AIDS. URL: https://www.unaids.org/en/resources/presscentre/featurestories/2020/october/20201015_chatbot [accessed 2023-08-15]
30. Garett R, Young SD. Potential application of conversational agents in HIV testing uptake among high-risk populations. J Public Health (Oxf) 2023 Mar 14;45(1):189-192 [FREE Full text] [doi: 10.1093/pubmed/fdac020] [Medline: 35211740]
31. Prakash AV, Das S. Intelligent conversational agents in mental healthcare services: a thematic analysis of user perceptions. Pac Asia J Assoc Inf Syst 2020 Jun 30;12:1-34 [FREE Full text] [doi: 10.17705/1thci.12201]

Abbreviations

AI: artificial intelligence
mHealth: mobile health
MSM: men who have sex with men
PrEP: pre-exposure prophylaxis
UTAUT: Unified Theory of Acceptance and Use of Technology

Edited by A Kushniruk, E Borycki; submitted 21.08.23; peer-reviewed by S Comulada, S John, R Saifi; comments to author 19.09.23; revised version received 22.09.23; accepted 12.12.23; published 26.01.24.

Please cite as:
Cheah MH, Gan YN, Altice FL, Wickersham JA, Shrestha R, Salleh NAM, Ng KS, Azwa I, Balakrishnan V, Kamarulzaman A, Ni Z Testing the Feasibility and Acceptability of Using an Artificial Intelligence Chatbot to Promote HIV Testing and Pre-Exposure Prophylaxis in Malaysia: Mixed Methods Study
JMIR Hum Factors 2024;11:e52055
URL: https://humanfactors.jmir.org/2024/1/e52055
doi:10.2196/52055
PMID:38277206

©Min Hui Cheah, Yan Nee Gan, Frederick L Altice, Jeffrey A Wickersham, Roman Shrestha, Nur Afiqah Mohd Salleh, Kee Seong Ng, Iskandar Azwa, Vimala Balakrishnan, Adeeba Kamarulzaman, Zhao Ni. Originally published in JMIR Human Factors (https://humanfactors.jmir.org), 26.01.2024. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic

XSL•FO
RenderX

information, a link to the original publication on https://humanfactors.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Viewpoint

The Temperature Feature of ChatGPT: Modifying Creativity for Clinical Research

Joshua Davis[1,2*], BS; Liesbet Van Bulck[3,4*], RN, PhD; Brigitte N Durieux[1], BS; Charlotta Lindvall[1,5,6], MD, PhD

[1]Department of Psychosocial Oncology and Palliative Care, Dana-Farber Cancer Institute, Boston, MA, United States
[2]Albany Medical College, Albany, NY, United States
[3]KU Leuven Department of Public Health and Primary Care, KU Leuven-University of Leuven, Leuven, Belgium
[4]Research Foundation Flanders (FWO), Brussels, Belgium
[5]Department of Medicine, Brigham and Women's Hospital, Boston, MA, United States
[6]Harvard Medical School, Harvard University, Boston, MA, United States
[*]these authors contributed equally

Corresponding Author:
Charlotta Lindvall, MD, PhD
Department of Psychosocial Oncology and Palliative Care, Dana-Farber Cancer Institute
450 Brookline Ave
Boston, MA, 02215
United States
Phone: 1 617 632 6464
Email: charlotta_lindvall@dfci.harvard.edu

Abstract

More clinicians and researchers are exploring uses for large language model chatbots, such as ChatGPT, for research, dissemination, and educational purposes. Therefore, it becomes increasingly relevant to consider the full potential of this tool, including the special features that are currently available through the application programming interface. One of these features is a variable called temperature, which changes the degree to which randomness is involved in the model's generated output. This is of particular interest to clinicians and researchers. By lowering this variable, one can generate more consistent outputs; by increasing it, one can receive more creative responses. For clinicians and researchers who are exploring these tools for a variety of tasks, the ability to tailor outputs to be less creative may be beneficial for work that demands consistency. Additionally, access to more creative text generation may enable scientific authors to describe their research in more general language and potentially connect with a broader public through social media. In this viewpoint, we present the temperature feature, discuss potential uses, and provide some examples.

(JMIR Hum Factors 2024;11:e53559) doi:10.2196/53559

KEYWORDS

artificial intelligence; ChatGPT; clinical communication; creative; creativity; customization; customize; customized; generation; generative; language model; language models; LLM; LLMs; natural language processing; NLP; random; randomness; tailor; tailored; temperature; text; texts; textual

Introduction

ChatGPT [1] is a large language model developed by OpenAI that currently has over 100 million users [2]. As its popularity continues to grow, clinicians and researchers are among many considering its potential applications in health care and academia. In a short time, ChatGPT has been extensively published [3], with clinical researchers exploring its potential utility for a variety of tasks, including answering patient questions [4,5], generating clinical summaries [6], and abstracting data from important documentation (eg, computed tomography reports) [7].

When using ChatGPT, one can interact through the website by providing a single prompt or engaging in a conversation. In addition to this more well-known web-based version of ChatGPT, there is also an application programming interface (API) that allows for more customization and flexibility. With the API, users can programmatically interact with ChatGPT and modify features for their specific use case. Although this approach may currently require more technical expertise for clinicians to use, its features may become available on the web

XSL•FO
RenderX

interface in future iterations of the tool. Therefore, these features are important to understand and relevant to discuss in terms of their meaning for clinicians and researchers in advance of their more widespread use. Additionally, they have direct implications for introducing greater reproducibility in use cases where this matters.

The Temperature Feature of ChatGPT

ChatGPT generates text through a probabilistic language modeling approach, where it writes responses word by word, calculating the most likely next word in the sequence. A key feature that influences this behavior is called temperature [8, 9]. In this context, temperature is a value from 0 to 2 that adjusts how random each subsequent word in the chat output is. A value of 0 will give the most probable word and, thus, the least variability. As the value increases toward and beyond 1, the next word becomes less probable, leading to more randomness and "creativity" in the response. This feature can currently be adjusted in the API, where the default value is 1 [9].

The ability to adjust the "creativity" of ChatGPT output should also be of interest to clinicians and medical researchers using the tool. By accounting for temperature, large language models such as ChatGPT can be tailored for different use cases. Lowering ChatGPT's creativity level would be preferable for tasks that require more consistent outputs; for clinicians and researchers, tasks of this sort may include summarizing patient data (eg, symptoms and medications) or streamlining administrative tasks (eg, billing inquiries and patient registrations). Alternatively, increasing the creativity level may provide better outputs for less structured tasks and may specifically hold relevance for improving clinicians' and researchers' digital communication with other health care workers, patients, and a wider audience.

Currently, many clinicians may find it difficult to engage with the broader public when it comes to disseminating findings. A lack of public accessibility can limit the impact of research. Creativity proves to be a powerful ally in communication, as it simplifies complex concepts, fosters critical thinking, and enhances the accessibility of intricate information. These factors collectively contribute to increased audience engagement and more effective message transmission. This feature has value for developing patient education materials, medical courses, science communication, and public research dissemination [10]. Increasing the creativity level could be useful when trying to report research results on social media [11], at conferences, or in scientific publications. Among other things, researchers could use a more creative ChatGPT to help come up with tags that apply to their work, brainstorm analogies or metaphors for scientific concepts, and check the tone of their writing.

Illustrating This Feature: Example

As an illustration of the "creativity" feature in ChatGPT, we have provided ChatGPT with the abstract of a recently published article on the development of a keyword library for capturing symptom talk in oncology [12]. While setting the temperature to low (0.0), medium (0.5), and high (1.0), we asked ChatGPT to create a tweet for a public audience, a title for a scientific journal article, and a title for a keynote address. Example ChatGPT-generated responses can be found in Table 1.

Table 1. ChatGPT-generated responses on the prompts for the different temperatures. For each temperature level, the same prompts were run 10 times in a ChatGPT model (text-davinci-003).

Temperature	Low (0.0)	Medium (0.5)	High (1.0)
Tweet for public audience	We developed a keyword library to detect symptom talk in transcribed clinical conversations. Our findings suggest that physical symptoms are more easily detected than psychological symptoms, and that more advanced computational models are needed to detect symptom expressions. #AI #Healthcare #SymptomDetection	Have you heard about the new study that developed and evaluated a natural language processing keyword library to detect symptom talk in oncology conversations? The findings highlight the need for more advanced computational models to detect symptom expressions in transcribed clinical conversations. #Oncology #SymptomTalk #NLP #ComputationalMethods	Check out how computational methods can help us detect patient symptoms in clinical conversations! A recent study developed & evaluated a keyword library for capturing expressed symptoms, highlighting the need for more advanced models for real-time detection of symptoms at scale. #AI @healthinformatics
Title for scientific journal	Detecting Symptom Talk in Transcribed Outpatient Oncology Conversations: Developing and Evaluating a Natural Language Processing Keyword Library	Keyword-Based Detection of Symptom Talk in Oncology Conversations: Developing a Natural Language Processing Library and Exploring Opportunities for Advanced Computational Models	Unlocking the Ambiguity of Symptom Talk: Developing and Evaluating a Natural Language Processing Library for Detecting Outpatient Oncology Conversations
Title of a keynote address	Unlocking the Potential of Computational Methods for Detecting Symptoms: Achieving Better Patient Care Through Natural Language Processing	Harnessing the Power of Computational Methods to Detect Patient Suffering: A Study of Symptom Talk in Oncology Conversations	The Ambiguity of Symptom Expressions: Utilizing Computational Methods to Better Attend to Patient Suffering

The examples shown in Table 1 illustrate that when the creativity level in ChatGPT is adjusted, slightly different responses are generated; these different creativity levels may provide more useful output depending on the task at hand. For example, a tweet created with a high level of creativity includes an exclamation mark and directly addresses the audience. Therefore, it may be more engaging compared to tweets with a low or medium creativity level. The title of the article and keynote generated with high "creativity" were more surprising and potentially less useful for these tasks, although this may

XSL•FO
RenderX

depend on the context, setting, and personality of the user. For these tasks, the low- and medium-creativity titles were more straightforward. Importantly, these lower values do translate to more consistent responses. We ran each of these prompts 10 times, and at a temperature level of 0, all responses were identical. Given ChatGPT's normally variable output, this feature holds exciting implications for scenarios where consistency and reproducibility are preferred.

In addition to the results reported above, we have also experimented with adjustments in temperature level using other ChatGPT models (ie, gpt-3.5-turbo-1106, gpt-3.5-turbo-instruct, and gpt-4-1106-preview). All outputs appear in Multimedia Appendix 1**.** In contrast to what we found when using the ChatGPT model "text-davinci-003," some other models showed some variability, even at a temperature level of 0. Regardless, the relative variability of outputs is still modified by temperature, with a higher temperature increasing creativity. Users should consider and test how temperature impacts outputs within the model they are using.

In the examples provided above, we have demonstrated how adjusting the level of creativity can enhance science communication, making it more engaging. However, it is crucial to also acknowledge the potential risks associated with increasing creativity, especially for clinical cases. Using ChatGPT with high creativity settings in clinical contexts, such as for summarizing patient medical data, can be problematic. Excessive creativity might lead to the embellishment or misrepresentation of crucial information, either by omitting vital details or interpreting data too liberally. Such inaccuracies could impact patient treatment and outcomes. Therefore, it is advisable to lower the creativity level of ChatGPT in clinical applications. By doing so, we ensure that the summarized information remains faithful to the original data, thereby prioritizing accuracy and reliability over creative expression.

In summary, the temperature feature of ChatGPT allows users to adjust the level of "creativity." Although no previous articles have discussed or investigated this feature for its use in clinical research, it shows promising potential for clinicians and researchers. Both high and low creativity levels could have interesting applications for health care and may broaden the ways clinicians and researchers consider using artificial intelligence (AI) tools to close gaps in areas such as digital communication. ChatGPT documentation suggests using a temperature value of 0 to 0.2 for more focused (less creative) tasks and 0.8 to 1 for more random (more creative) tasks [9]. As large language models are variable and use case dependent, we strongly suggest testing and validating the proper temperature level for your specific use case. While this feature is a powerful tool that could be useful for creating easy-to-understand summaries, captivating social media posts, or making complex information more accessible to a wider audience, the parameters need to be carefully tweaked to find a balance between coherence and creativity and to tailor to specific needs. Looking ahead, as AI continues to advance in the health care sector, the temperature feature can play a pivotal role in health care applications in generative AI, unlocking the potential for more accurate, empathetic, or creative interactions between AI and health care stakeholders.

Acknowledgments

This work was supported by the Research Foundation Flanders (FWO; grants 1159522N and K210723N to LVB) and by the National Institutes of Health, National Institute on Aging Medical Student Training in Aging Research Program (grant 5T35AG038027-13 to JD). The funding sources had no involvement in the study design, analyses, or interpretation of the data; writing of the report; or decision to submit the paper for publication. During the preparation of this work, the authors used ChatGPT in order to generate the examples reported in the paper and to assist with the writing process. The use of a large language model has greatly assisted us in rephrasing and ensuring the clarity and effectiveness of our language. After using this tool, the authors reviewed and edited the content as needed and take full responsibility for the content of the publication.

Data Availability

All data generated or analyzed during this study are included in this published article and Multimedia Appendix 1.

Authors' Contributions

JD and LVB contributed equally and share first authorship. LVB, JD, and BND contributed to the conception and design of the study and drafted the paper. CL critically revised the paper. JD and LVB both accessed and verified the underlying data reported in the manuscript. All authors approved the final version of the manuscript and had full responsibility for the decision to submit for publication.

Conflicts of Interest

None declared.

Multimedia Appendix 1

All data presented in this article, ChatGPT outputs for tests of 3 prompts across 3 temperature values for 3 different models (gpt-3.5-turbo-1106, gpt-3.5-turbo-instruct, and gpt-4-1106-preview; 100 runs for each test), and a summary document describing the multiple model tests.

XSL•FO
RenderX

[ZIP File (Zip Archive), 241 KB - humanfactors_v11i1e53559_app1.zip]

References

1. ChatGPT. OpenAI. URL: https://openai.com/chatgpt [accessed 2024-02-16]
2. Hu K. ChatGPT sets record for fastest-growing user base—analyst note. Reuters. 2023. URL: https://www.reuters.com/technology/chatgpt-sets-record-fastest-growing-user-base-analyst-note-2023-02-01/ [accessed 2023-10-05]
3. Ward E, Gross C. Evolving methods to assess chatbot performance in health sciences research. JAMA Intern Med 2023;183(9):1030-1031. [doi: 10.1001/jamainternmed.2023.2567] [Medline: 37459112]
4. Ayers JW, Poliak A, Dredze M, Leas EC, Zhu Z, Kelley JB, et al. Comparing physician and artificial intelligence chatbot responses to patient questions posted to a public social media forum. JAMA Intern Med 2023;183(6):589-596. [doi: 10.1001/jamainternmed.2023.1838] [Medline: 37115527]
5. Van Bulck L, Moons P. What if your patient switches from Dr. Google to Dr. ChatGPT? a vignette-based survey of the trustworthiness, value, and danger of ChatGPT-generated responses to health questions. Eur J Cardiovasc Nurs 2024;23(1):95-98 [FREE Full text] [doi: 10.1093/eurjcn/zvad038] [Medline: 37094282]
6. Nayak A, Alkaitis MS, Nayak K, Nikolov M, Weinfurt KP, Schulman K. Comparison of history of present illness summaries generated by a chatbot and senior internal medicine residents. JAMA Intern Med 2023;183(9):1026-1027. [doi: 10.1001/jamainternmed.2023.2561] [Medline: 37459091]
7. Fink MA, Bischoff A, Fink CA, Moll M, Kroschke J, Dulz L, et al. Potential of ChatGPT and GPT-4 for data mining of free-text CT reports on lung cancer. Radiology 2023;308(3):e231362. [doi: 10.1148/radiol.231362] [Medline: 37724963]
8. Best practices for prompt engineering with OpenAI API: how to give clear and effective instructions to GPT-3 and Codex. Shieh J. OpenAI. 2023. URL: https://help.openai.com/en/articles/6654000-best-practices-for-prompt-engineering-with-openai-api [accessed 2023-12-11]
9. API reference. OpenAI. 2023. URL: https://platform.openai.com/docs/api-reference/chat [accessed 2023-12-11]
10. Heng JJY, Teo DB, Tan LF. The impact of Chat Generative Pre-Trained Transformer (ChatGPT) on medical education. Postgrad Med J 2023;99(1176):1125-1127. [doi: 10.1093/postmj/qgad058] [Medline: 37466157]
11. Grajales FJ, Sheps S, Ho K, Novak-Lauscher H, Eysenbach G. Social media: a review and tutorial of applications in medicine and health care. J Med Internet Res 2014;16(2):e13 [FREE Full text] [doi: 10.2196/jmir.2912] [Medline: 24518354]
12. Durieux BN, Zverev SR, Tarbi EC, Kwok A, Sciacca K, Pollak KI, et al. Development of a keyword library for capturing PRO-CTCAE-focused "symptom talk" in oncology conversations. JAMIA Open 2023;6(1):ooad009 [FREE Full text] [doi: 10.1093/jamiaopen/ooad009] [Medline: 36789287]

Abbreviations

AI: artificial intelligence
API: application programming interface

Edited by E Borycki, A Kushniruk; submitted 11.10.23; peer-reviewed by L Chinthala, C Li, S Minshall; comments to author 20.11.23; revised version received 11.12.23; accepted 24.01.24; published 08.03.24.

Please cite as:
Davis J, Van Bulck L, Durieux BN, Lindvall C
The Temperature Feature of ChatGPT: Modifying Creativity for Clinical Research
JMIR Hum Factors 2024;11:e53559
URL: https://humanfactors.jmir.org/2024/1/e53559
doi:10.2196/53559
PMID:38457221

©Joshua Davis, Liesbet Van Bulck, Brigitte N Durieux, Charlotta Lindvall. Originally published in JMIR Human Factors (https://humanfactors.jmir.org), 08.03.2024. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on https://humanfactors.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

The Impact of Performance Expectancy, Workload, Risk, and Satisfaction on Trust in ChatGPT: Cross-Sectional Survey Analysis

Avishek Choudhury[1], PhD; Hamid Shamszare[1], MSc

Industrial and Management Systems Engineering, Benjamin M. Statler College of Engineering and Mineral Resources, West Virginia University, Morgantown, WV, United States

Corresponding Author:
Avishek Choudhury, PhD
Industrial and Management Systems Engineering
Benjamin M. Statler College of Engineering and Mineral Resources
West Virginia University
321 Engineering Sciences Building
1306 Evansdale Drive
Morgantown, WV, 26506
United States
Phone: 1 3042939431
Email: avishek.choudhury@mail.wvu.edu

Abstract

Background: ChatGPT (OpenAI) is a powerful tool for a wide range of tasks, from entertainment and creativity to health care queries. There are potential risks and benefits associated with this technology. In the discourse concerning the deployment of ChatGPT and similar large language models, it is sensible to recommend their use primarily for tasks a human user can execute accurately. As we transition into the subsequent phase of ChatGPT deployment, establishing realistic performance expectations and understanding users' perceptions of risk associated with its use are crucial in determining the successful integration of this artificial intelligence (AI) technology.

Objective: The aim of the study is to explore how perceived workload, satisfaction, performance expectancy, and risk-benefit perception influence users' trust in ChatGPT.

Methods: A semistructured, web-based survey was conducted with 607 adults in the United States who actively use ChatGPT. The survey questions were adapted from constructs used in various models and theories such as the technology acceptance model, the theory of planned behavior, the unified theory of acceptance and use of technology, and research on trust and security in digital environments. To test our hypotheses and structural model, we used the partial least squares structural equation modeling method, a widely used approach for multivariate analysis.

Results: A total of 607 people responded to our survey. A significant portion of the participants held at least a high school diploma (n=204, 33.6%), and the majority had a bachelor's degree (n=262, 43.1%). The primary motivations for participants to use ChatGPT were for acquiring information (n=219, 36.1%), amusement (n=203, 33.4%), and addressing problems (n=135, 22.2%). Some participants used it for health-related inquiries (n=44, 7.2%), while a few others (n=6, 1%) used it for miscellaneous activities such as brainstorming, grammar verification, and blog content creation. Our model explained 64.6% of the variance in trust. Our analysis indicated a significant relationship between (1) workload and satisfaction, (2) trust and satisfaction, (3) performance expectations and trust, and (4) risk-benefit perception and trust.

Conclusions: The findings underscore the importance of ensuring user-friendly design and functionality in AI-based applications to reduce workload and enhance user satisfaction, thereby increasing user trust. Future research should further explore the relationship between risk-benefit perception and trust in the context of AI chatbots.

(JMIR Hum Factors 2024;11:e55399) doi:10.2196/55399

KEYWORDS

ChatGPT; chatbots; health care; health care decision-making; health-related decision-making; health care management; decision-making; user perception; usability; usable; usableness; usefulness; artificial intelligence; algorithms; predictive models; predictive analytics; predictive system; practical models; deep learning; cross-sectional survey

XSL•FO
RenderX

Introduction

ChatGPT (OpenAI) [1] is a powerful tool for a wide range of tasks, from entertainment and creativity to health care queries [2]. However, there are potential benefits associated with this technology. For instance, it can help summarize large amounts of text data [3,4] or generate programming code [5]. There is also the notion that ChatGPT may potentially assist with health care tasks [6-9]. However, the risks associated with using ChatGPT can hinder its adoption in various high-risk domains. These risks include the potential for inaccuracies and lack of citation relevance in scientific content generated by ChatGPT [10], ethical issues (copyright, attribution, plagiarism, and authorship) [11], the risk of hallucination (inaccurate information that sounds plausible scientifically) [12], and the possibility of biased content and inaccurate information due to the quality of training data sets generated prior to the year 2021 [4].

In the discourse concerning the deployment of ChatGPT and similar artificial intelligence (AI) technologies, it is sensible to recommend their use primarily for tasks a human user can execute accurately. Few studies have advocated using the technology under human supervision [13,14]. Encouraging users to rely on such tools for tasks beyond their competence is risky, as they may need help to evaluate the AI's output effectively. The strength of ChatGPT lies in its ability to automate more straightforward, mundane tasks, freeing human users to invest their time and cognitive resources into critical tasks (not vice versa). This approach to technology use maintains a necessary balance, leveraging AI for efficiency gains while ensuring that critical decision-making remains within the purview of human expertise.

As we transition into the subsequent phase of ChatGPT deployment, establishing realistic performance expectations and understanding users' perceptions of risk associated with its use are crucial in determining the successful integration of this AI technology. Thus, understanding users' perceptions of ChatGPT becomes essential, as these perceptions significantly influence their usage decisions [2]. For example, suppose users believe that ChatGPT's capabilities surpass human knowledge. In that case, they may be tempted to use it for tasks such as self-diagnosis, which could lead to potentially harmful outcomes if the generated information is mistaken or misleading. Conversely, a realistic appraisal of the limitations and strengths of technology would encourage its use in low-risk, routine tasks and foster a safer, more effective integration into our everyday lives.

Building upon the importance of user perceptions and expectations, we must also consider that the extent to which users trust ChatGPT hinges mainly on the perception of its accuracy and reliability. As users witness the technology's ability to perform tasks effectively and generate correct, helpful information, their trust in the system grows. This, in turn, allows them to offload routine tasks to the AI and focus their energies on more complex or meaningful endeavors. Similarly, instances where the AI generates inaccurate or misleading information can quickly erode users' perception of the technology. Users may become dissatisfied and lose trust if they perceive the technology as unreliable or potentially harmful, particularly if they have previously overestimated its capabilities. This underlines the importance of setting realistic expectations and accurately understanding the strengths and limitations of ChatGPT, which can help foster a healthy level of trust and satisfaction among users. Ultimately, establishing and maintaining trust and satisfaction are not a onetime event but an ongoing process of validating the AI's outputs, understanding and acknowledging its limitations, and making the best use of its capabilities within a framework of informed expectations and continuous learning. This dynamic balance is pivotal for the effective and safe integration of AI technologies such as ChatGPT into various sectors of human activity.

In our prior work, we explored the impact of trust in the actual use of ChatGPT [15]. This study aims to explore a conceptual framework delving deeper into the aspects influencing user trust in ChatGPT.

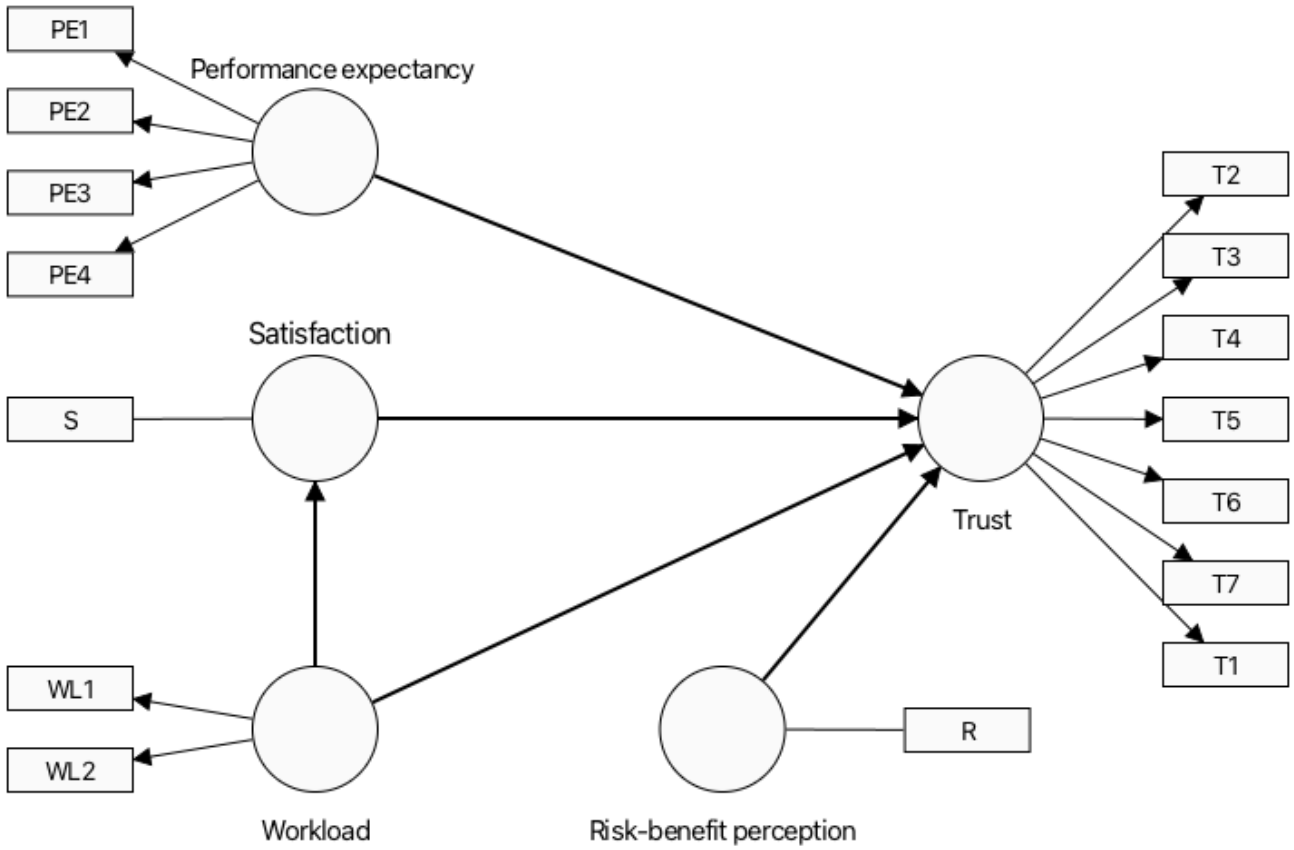
As shown in Figure 1, the proposed conceptual model is grounded in the well-established theories of technology acceptance and use, incorporating constructs such as performance expectancy, workload, satisfaction, risk-benefit perception, and trust to comprehensively evaluate user interaction with technology. Performance expectancy, derived from the core postulates of the technology acceptance model (TAM) [16] and the unified theory of acceptance and use of technology (UTAUT) [17], posits that the perceived use of the technology significantly predicts usage intentions. Workload, akin to effort expectancy, reflects the perceived cognitive and physical effort required to use the technology, where a higher workload may inversely affect user satisfaction—a construct that encapsulates the fulfillment of user expectations and needs through technology interaction. The risk-benefit perception embodies the user's assessment of the technology's potential advantages against its risks, intricately influencing both user satisfaction and trust. Trust, a pivotal determinant of technology acceptance [15], signifies the user's confidence in the reliability and efficacy of the technology. This theoretical framework thus serves to elucidate the multifaceted process by which users come to accept and use a technological system, highlighting the critical role of both cognitive appraisals and affective responses in shaping the technology adoption landscape.

We explore the following hypotheses:

- Hypothesis 1: Perceived workload of using ChatGPT negatively correlates with user trust in ChatGPT.
- Hypothesis 2: Perceived workload of using ChatGPT negatively correlates with user satisfaction with ChatGPT.
- Hypothesis 3: User satisfaction with ChatGPT positively correlates with trust in ChatGPT.
- Hypothesis 4: User trust in ChatGPT is positively correlated with the performance expectancy of ChatGPT.
- Hypothesis 5: The risk-benefit perception of using ChatGPT is positively correlated with user trust in ChatGPT.

XSL•FO
RenderX

Figure 1. A conceptual model of technology acceptance illustrating trust (T) as the dependent outcome variable, with performance expectancy (PE), workload (WL), and risk-benefit perception (R) as direct predictors. Satisfaction (S) is depicted as a mediating variable that moderates the impact of workload on trust.



Methods

Ethical Considerations

The study obtained ethics approval from West Virginia University, Morgantown (protocol 2302725983). The study was performed in accordance with relevant guidelines and regulations. No identifiers were collected during the study, and all users were compensated for completing the survey through an audience paneling service. In compliance with ethical research practices, informed consent was obtained from all participants before initiating the survey. Attached to the survey was a comprehensive cover letter outlining the purpose of the study, the procedure involved, the approximate time to complete the survey, and assurances of anonymity and confidentiality. It also emphasized that participation was completely voluntary, and participants could withdraw at any time without any consequences. The cover letter also included the contact information of the researchers for any questions or concerns the participants might have regarding the study. Participants were asked to read through the cover letter information carefully and were instructed to proceed with the survey only if they understood and agreed to the terms described, effectively providing their consent to participate in the study.

Study Design

A semistructured, web-based questionnaire was disseminated to adult individuals within the United States who engaged with ChatGPT (version 3.5) at least once per month. Data collection took place between February and March 2023. The questionnaire was crafted using Qualtrics (Qualtrics LLC), and its circulation was handled by Centiment (Centiment LLC), a provider of audience-paneling services. Centiment's services were used due to their extensive reach and ability to connect with a diverse and representative group via their network and social media. Their fingerprinting technology, which uses IP address, device type, screen size, and cookies, was used to guarantee the uniqueness of the survey respondents. Prior to the full-scale dissemination, a soft launch was carried out with 40 responses gathered. The purpose of a soft launch, a limited-scale trial of the survey, is to pinpoint any potential problems, such as ambiguity or confusion in questions, technical mishaps, or any other factors that might affect the quality of data obtained. The survey was made available to a larger audience following the successful soft launch.

Table 1 shows the descriptive statistics of the survey questions used in this study. We developed 3 latent constructs based on the question: trust, workload, and performance expectancy, and 2 single question variables: satisfaction and risk-benefit perception. Participant responses to all the questions were captured using a 4-point Likert scale ranging from 1=strongly disagree to 4=strongly agree. These questions were adapted from constructs used in various models and theories such as the TAM, the theory of planned behavior, UTAUT, and research on trust and security in digital environments.

- Trust: Questions T1-T7 related to trust in AI systems were adapted from the trust building model [18].

XSL•FO
RenderX

- Workload: WL1 and WL2 questions from the National Aeronautics and Space Administration Task Load Index for measuring perceived workload [19].
- Performance expectancy: PE1-PE4 are about the perceived benefits of using the system, which is a central concept in TAM and UTAUT.
- Satisfaction: The single item relates to overall user satisfaction, a common measure in information systems success models [20].
- Risk-benefit perception: Question addresses the user's assessment of benefits relative to potential risks, an aspect often discussed in the context of technology adoption and use [21].

These references provide a starting point for understanding the theoretical underpinnings of the survey used in this study. They are adapted from foundational works in information systems, human-computer interaction, and psychology that address trust, workload, performance expectancy, satisfaction, and the evaluation of benefits versus risks in technology use.

- Trust: Questions T1-T7 related to trust in AI systems were adapted from the trust building model [18].
- Workload: WL1 and WL2 questions from the National Aeronautics and Space Administration Task Load Index for measuring perceived workload [19].
- Performance expectancy: PE1-PE4 are about the perceived benefits of using the system, which is a central concept in TAM and UTAUT.
- Satisfaction: The single item relates to overall user satisfaction, a common measure in information systems success models [20].
- Risk-benefit perception: Question addresses the user's assessment of benefits relative to potential risks, an aspect often discussed in the context of technology adoption and use [21].

These references provide a starting point for understanding the theoretical underpinnings of the survey used in this study. They are adapted from foundational works in information systems, human-computer interaction, and psychology that address trust, workload, performance expectancy, satisfaction, and the evaluation of benefits versus risks in technology use.

Table 1. Study variables and latent construct (N=607).

Survey items	Value, mean (SD)
Trust (T)	
T1: ChatGPT is competent in providing the information and guidance I need	3.20 (0.83)
T2: ChatGPT is reliable in providing consistent and dependable information	3.16 (0.80)
T3: ChatGPT is transparent	3.12 (0.86)
T4: ChatGPT is trustworthy in the sense that it is dependable and credible	3.17 (0.84)
T5: ChatGPT will not cause harm, manipulate its responses, or create negative consequences for me	3.10 (0.88)
T6: ChatGPT will act with integrity and be honest with me	3.19 (0.82)
T7: ChatGPT is secure and protects my privacy and confidential information	3.27 (0.81)
Workload (WL)	
WL1: Using ChatGPT was mentally demanding	3.21 (0.75)
WL2: I had to work hard to use ChatGPT	2.20 (0.98)
Performance expectancy (PE)	
PE1: ChatGPT can help me achieve my goals	3.24 (0.77)
PE2: ChatGPT can reduce my workload	3.22 (0.78)
PE3: ChatGPT improves my work efficiency	3.21 (0.84)
PE4: ChatGPT helps me make informed and timely decisions	3.26 (0.79)
Satisfaction (S)	
S: I am satisfied with ChatGPT	3.24 (0.76)
Risk-benefit perception (R)	
R: The benefits of using ChatGPT outweigh any potential risks	3.20 (0.80)

Statistical Analysis and Model Validation

To test our hypotheses and structural model, we used the partial least squares structural equation modeling (PLS-SEM) method, a widely used approach for multivariate analysis. PLS-SEM enables the estimation of complex models with multiple constructs, indicator variables, and structural paths, without making assumptions about the data's distribution [22]. This method is beneficial for studies with small sample sizes that involve many constructs and items [23]. PLS-SEM is a suitable method because of its flexibility and ability to allow for interaction between theory and data in exploratory research [24]. The analyses were performed using the *SEMinR* package

XSL•FO
RenderX

in R (R Foundation for Statistical Computing) [25]. We started by loading the data set collected for this study using the *reader* package in R. We then defined the measurement model. This consisted of 5 composite constructs: trust, performance expectancy, workload, risk-benefit perception, and satisfaction. Trust was measured with 7 items (T1 through T7), performance expectancy with 4 items (PE1 through PE4), and workload with 2 items (WL1 and WL2), while risk-benefit perception and satisfaction were each measured with a single item. We also evaluated the convergent and discriminant validity of the latent constructs, which we assessed using 3 criteria: factor loadings (>0.50), composite reliability (>0.70), and average variance extracted (>0.50). We used the Heterotrait-Monotrait ratio (<0.90) to assess discriminant validity [26].

Next, we defined the structural model, which captured the hypothesized relationships between the constructs. The model included paths from risk-benefit perception, performance expectancy, workload, satisfaction to trust, and a path from workload to satisfaction. We then estimated the model's parameters using the partial least squares method. This was done with the *estimate_pls* function in the *seminar* package. The partial least squares method was preferred due to its ability to handle complex models and its robustness to violations of normality assumptions. We performed a bootstrap resampling procedure with 10,000 iterations to obtain robust parameter estimates and compute 95% CIs. The bootstrapped model was plotted to visualize the estimates and their 95% CIs.

Results

Of 607 participants who completed the survey, 29.9% (n=182) used ChatGPT at least once per month, 26.1% (n=158) used it weekly, 24.5% (n=149) accessed it more than once per week, and 19.4% (n=118) interacted with it almost daily. A substantial portion of the participants held at least a high school diploma (n=204, 33.6%), and the majority had a bachelor's degree (n=262, 43.1%). The primary motivations for participants to use ChatGPT were for acquiring information (n=219, 36%), amusement (n=203, 33.4%), and addressing problems (n=135, 22.2%). Some participants used it for health-related inquiries (n=44, 7.2%), while a few others (n=6, 1%) used it for miscellaneous activities such as brainstorming, grammar verification, and blog content creation. Table 2 shows the factor loading of the latent constructs in the model.

The model explained 2% and 64.6% of the variance in "satisfaction" and "trust," respectively. Reliability estimates, as shown in Table 3, indicated high levels of internal consistency for all 5 latent variables, with Cronbach α and ρ values exceeding the recommended threshold of 0.7. The average variance extracted for the latent variables also exceeded the recommended threshold of 0.5, indicating that these variables are well-defined and reliable. Based on the root mean square error of approximation (RMSEA) fit index, our PLS-SEM model demonstrates a good fit for the observed data. The calculated RMSEA value of 0.07 falls below the commonly accepted threshold of 0.08, indicating an acceptable fit. The RMSEA estimates the average discrepancy per degree of freedom in the model, capturing how the proposed model aligns with the population covariance matrix. With a value below the threshold, it suggests that the proposed model adequately represents the relationships among the latent variables. This finding provides confidence in the model's ability to explain the observed data and support the underlying theoretical framework.

Table 4 shows the estimated paths in our model. Hypothesis 1 postulated that as the perceived workload of using ChatGPT increases, user trust in ChatGPT decreases. Our analysis indicated a negative estimate for the path from workload to trust (–0.047). However, the *T* statistic (–1.674) is less than the critical value, and the 95% CI straddles 0 (–0.102 to –0.007), suggesting that the effect is not statistically significant. Therefore, we do not have sufficient evidence to support hypothesis 1.

Hypothesis 2 stated that perceived workload is negatively correlated with user satisfaction with ChatGPT. The results supported this hypothesis, as the path from workload to satisfaction showed a negative estimate (–0.142), a *T* statistic (–3.416) beyond the critical value, and a 95% CI (–0.223 to –0.061).

The data confirmed this relationship for hypothesis 3, which proposed a positive correlation between satisfaction with ChatGPT and trust in ChatGPT. The path from satisfaction to trust had a positive estimate (0.165), a *T* statistic (4.478) beyond the critical value, and a 95% CI (0.093-0.237).

Hypothesis 4 suggested that user performance expectations of ChatGPT increase with their trust in the technology. The analysis supported this hypothesis. The path from performance expectancy to trust displayed a positive estimate (0.598), a large *T* statistic (15.554), and a 95% CI (0.522-0.672). Finally, we examined hypothesis 5, which posited that user trust in ChatGPT increases as their risk-benefit perception of using the technology increases. The path from risk-benefit perception to trust showed a positive estimate (0.114). The *T* statistic (3.372) and the 95% CI (0.048-0.179) indicating this relationship is significant, but the positive sign suggests that as the perceived benefits outweigh the risks, the trust in ChatGPT increases. Therefore, hypothesis 5 is supported. Figure 2 illustrates the structural model with all path coefficients.

XSL•FO
RenderX

Table 2. Bootstrapped loadings: model analysis estimates the relationship between various constructs and their indicators.

Bootstrapped loadings	Loadings	*T* statistic	95% CI
Trust (T)			
T1	0.788	41.998	0.750-0.823
T2	0.753	33.795	0.706-0.794
T3	0.773	40.293	0.733-0.808
T4	0.732	28.772	0.679-0.779
T5	0.673	21.066	0.607-0.732
T6	0.799	46.065	0.763-0.831
T7	0.779	38.088	0.736-0.816
Performance expectancy (PE)			
PE1	0.809	49.231	0.775-0.839
PE2	0.733	29.360	0.681-0.779
PE3	0.802	44.968	0.766-0.835
PE4	0.777	34.198	0.729-0.818
Workload (WL)			
WL1	0.856	28.883	0.789-0.905
WL2	0.913	44.872	0.869-0.950

Table 3. Convergent reliability.

Construct	Cronbach α	ρ C	AVE[a]	ρ A
Performance expectation	0.787	0.862	0.610	0.610
Workload	0.729	0.870	0.771	0.968
Trust	0.876	0.904	0.575	0.880

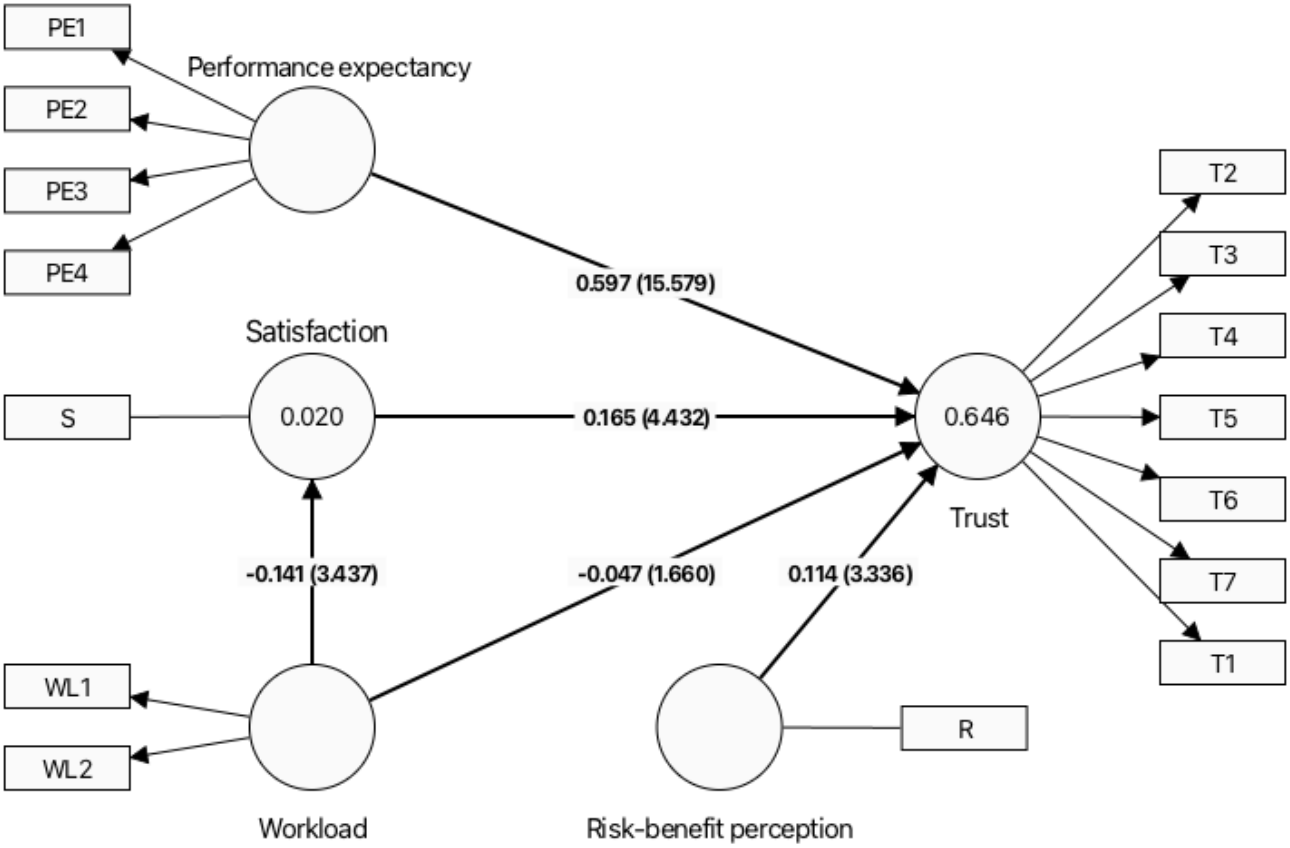
[a]AVE: average variance extracted.

Table 4. Bootstrapped structural path estimates.

Direct path	Bootstrap mean standard estimate (SD)	*T* statistic	95% CI
Risk-benefit perception→trust	0.114 (0.034)	3.372	0.048 to 0.179
Performance expectancy→trust	0.598 (0.038)	15.554	0.522 to 0.672
Workload→satisfaction	–0.142 (0.041)	–3.416	–0.223 to –0.061
Workload→trust	–0.047 (0.028)	–1.674	–0.102 to 0.007
Satisfaction→trust	0.165 (0.037)	4.478	0.093 to 0.237

XSL•FO
RenderX

Figure 2. The significant paths connecting trust (T) in ChatGPT, performance expectancy (PE), satisfaction (S), workload (WL), and risk-benefit perception (R). T1 through T7: factors for trust; PE1 through PE4: factors for performance expectancy; and WL1 and WL2: factors for workload. The inner model shows the path coefficient and T statistic values.



Discussion

Main Findings

This study represents one of the initial attempts to investigate how human factors such as workload, performance expectancy, risk-benefit perception, and satisfaction influence trust in ChatGPT. Our results showed that these factors significantly influenced trust in ChatGPT, with performance expectancy exerting the strongest association, highlighting its critical role in fostering trust. Additionally, we found that satisfaction was a mediator in the relationship between workload and trust. At the same time, a positive correlation was observed between trust in ChatGPT and the risk-benefit perception. Our findings align with the May 23, 2023, efforts and initiatives of the Biden-Harris Administration to advance responsible AI research, development, and deployment [27]. The Administration recognizes that managing its risks is crucial and prioritizes protecting individuals' rights and safety. One of the critical actions taken by the administration is the development of the artificial intelligence risk management framework (AI RMF). The AI RMF builds on the importance of trustworthiness in AI systems and is a framework for strengthening AI trustworthiness and promoting the trustworthy design, development, deployment, and use of AI systems, contributing to the need for our research [28]. Our findings reveal the importance of performance expectancy, satisfaction, and risk-benefit perception in determining the user's trust in AI systems. By addressing these factors, AI systems can be designed and developed to be more user-centric, aligning with the AI RMF's emphasis on human-centricity and responsible AI.

Workload and Trust in ChatGPT

Moreover, we found that reducing user workload is vital for enhancing user satisfaction, which in turn improves trust. This finding aligns with the AI RMF's focus on creating AI systems that are equitable and accountable and that mitigate inequitable outcomes. Additionally, our research emphasizes the need for future exploration of other factors impacting user trust in AI technologies. Such endeavors align with the AI RMF's vision of managing AI risks comprehensively and holistically, considering technical and societal factors. Understanding these factors is crucial for fostering public trust and enhancing the overall trustworthiness of AI systems, as outlined in the AI RMF [28].

This study also extends and complements existing literature. Consistent with the observed patterns in studies on flight simulators, dynamic multitasking environments, and cyberattacks [29-31], we also found that higher perceived workload in using ChatGPT led to lower levels of trust in this technology. Our findings align with the existing research indicating a negative correlation between workload and user satisfaction [32]. We observed that as the perceived workload of using ChatGPT increased, user satisfaction with the technology decreased. This outcome echoes the consensus within the literature that a high workload can lead to user dissatisfaction, particularly if the technology requires too much effort or time [33]. The literature reveals that perceived

XSL•FO
RenderX

workload balance significantly influences job satisfaction in work organizations [25], and similar patterns are found in the well-being studies of nurses, where perceived workload negatively impacts satisfaction with work-life balance [34]. While this study does not directly involve the workplace environment or work-life balance, the parallels between workload and satisfaction are evident. Furthermore, our research parallels the study suggesting that when providing timely service, AI applications can alleviate perceived workload and improve job satisfaction [35]. ChatGPT, as an AI-powered chatbot, could potentially contribute to workload relief when it performs effectively and efficiently, thereby boosting user satisfaction.

Satisfaction and Trust in ChatGPT

Our findings corroborate with existing literature, suggesting a strong positive correlation between user satisfaction and trust in the technology or service provider [23,24,26,36-38]. We found that the users who expressed higher satisfaction with ChatGPT were more likely to trust the system, strengthening the premise that satisfaction can predict trust in a technology or service provider. Similar to the study on digital transaction services, our research indicates that higher satisfaction levels with ChatGPT corresponded with higher trust in the AI system [37]. This suggests that when users are satisfied with the performance and results provided by ChatGPT, they tend to trust the technology more. The research on mobile transaction apps mirrors our findings, where we also discovered that satisfaction with ChatGPT use was a significant predictor of trust in the system [36]. This showcases the importance of ensuring user satisfaction in fostering trust using innovative technologies like AI chatbots. The study on satisfaction with using digital assistants, where a positive relationship between trust and satisfaction was observed [26], further aligns with our study. We also found a positive correlation between trust in ChatGPT and user satisfaction with this AI assistant.

Performance Expectancy and Trust in ChatGPT

Our findings concerning the strong positive correlation between performance expectancy and trust in ChatGPT serve as an extension to prior literature. Similar findings have been reported in previous studies on wearables and mobile banking [39,40], where performance expectancy was positively correlated with trust. However, our results diverge from the observations of a recent study that did not find a significant impact of performance expectancy on trust in chatbots [41]. Moreover, the observed mediating role of satisfaction in the relationship between workload and trust in ChatGPT is a notable contribution to the literature. While previous studies have demonstrated a positive correlation between workload reduction by chatbots and trust, as well as between trust and user satisfaction [42-44], the role of satisfaction as a mediator between workload and trust has not been explored. Finally, the positive correlation between the risk-benefit perception of using ChatGPT and trust aligns with the findings of previous studies [45-47]. Similar studies on the intention to use chatbots for digital shopping and customer service have found that trust in chatbots impacts perceived risk and is affected by the risk involved in using chatbots [46,47]. Our study adds to this body of research by confirming the same positive relationship within the context of ChatGPT.

Limitations

Despite the valuable insights provided by this study, limitations should be acknowledged. First, our research focused explicitly on ChatGPT and may not be generalizable to other AI-powered conversational agents or chatbot technologies. Different chatbot systems may have unique characteristics and user experiences that could influence the factors affecting trust. Second, this study relied on self-reported data from survey responses, which may be subject to response biases and limitations inherent to self-report measures. Participants' perceptions and interpretations of the constructs under investigation could vary, leading to potential measurement errors. Third, this study was cross-sectional, capturing data at a specific point in time. Longitudinal studies that track users' experiences and perceptions over time provide a more comprehensive understanding of the dynamics between trust and the factors investigated. Finally, the sample of participants in this study consisted of individuals who actively use ChatGPT, which may introduce a self-selection bias. The perspectives and experiences of nonusers or individuals with limited exposure to AI-powered conversational agents may differ, and their insights could provide additional valuable perspectives.

Conclusions

This study examined the factors influencing trust in ChatGPT, an AI-powered conversational agent. Our analysis found that performance expectancy, satisfaction, workload, and risk-benefit perceptions significantly influenced users' trust in ChatGPT. These findings contribute to understanding trust dynamics in the context of AI-powered conversational agents and provide insights into the factors that can enhance user trust. By addressing the factors influencing trust, we contribute to the broader goal of fostering responsible AI practices that prioritize user-centric design and protect individuals' rights and safety. Future research should consider longitudinal designs to capture the dynamics of trust over time. Additionally, incorporating perspectives from diverse user groups and examining the impact of contextual factors on trust would further enrich our understanding of trust in AI technologies.

Data Availability

The data sets generated and analyzed during this study are available from the corresponding author on reasonable request.

Authors' Contributions

AC, the lead researcher, was responsible for the study's conceptualization, the survey's development, figure illustration, data collection and analysis, and manuscript writing. HS, the student author, was responsible for manuscript writing and conducting

XSL•FO
RenderX

the literature review. Both authors collaborated throughout the research process and approved the final version of the manuscript for submission.

Conflicts of Interest

None declared.

References

1. OpenAI: Models GPT-3. OpenAI. URL: https://platform.openai.com/docs/models/gpt-4 [accessed 2024-05-07]
2. Shahsavar Y, Choudhury A. User intentions to use ChatGPT for self-diagnosis and health-related purposes: cross-sectional survey study. JMIR Hum Factors 2023;10:e47564 [FREE Full text] [doi: 10.2196/47564] [Medline: 37195756]
3. Editorial N. Tools such as ChatGPT threaten transparent science; here are our ground rules for their use. Nature 2023;613(7945):612 [FREE Full text] [doi: 10.1038/d41586-023-00191-1] [Medline: 36694020]
4. Moons P, Van Bulck L. ChatGPT: can artificial intelligence language models be of value for cardiovascular nurses and allied health professionals. Eur J Cardiovasc Nurs 2023;22(7):e55-e59 [FREE Full text] [doi: 10.1093/eurjcn/zvad022] [Medline: 36752788]
5. Aljanabi M, Ghazi M, Ali AH, Abed SA. ChatGpt: open possibilities. Iraqi J Sci Comput Sci Math 2023;4(1):62-64. [doi: 10.52866/20ijcsm.2023.01.01.0018]
6. D'Amico RS, White TG, Shah HA, Langer DJ. I asked a ChatGPT to write an editorial about how we can incorporate chatbots into neurosurgical research and patient care…. Neurosurgery 2023;92(4):663-664 [FREE Full text] [doi: 10.1227/neu.0000000000002414] [Medline: 36757199]
7. Holzinger A, Keiblinger K, Holub P, Zatloukal K, Müller H. AI for life: trends in artificial intelligence for biotechnology. N Biotechnol 2023;74:16-24 [FREE Full text] [doi: 10.1016/j.nbt.2023.02.001] [Medline: 36754147]
8. Sharma G, Thakur A. ChatGPT in drug discovery. ChemRxiv. 2023. URL: https://chemrxiv.org/engage/chemrxiv/article-details/63d56c13ae221ab9b240932f [accessed 2024-05-07]
9. Mann DL. Artificial intelligence discusses the role of artificial intelligence in translational medicine: a interview with ChatGPT. JACC Basic Transl Sci 2023;8(2):221-223 [FREE Full text] [doi: 10.1016/j.jacbts.2023.01.001] [Medline: 36908674]
10. Chen TJ. ChatGPT and other artificial intelligence applications speed up scientific writing. J Chin Med Assoc 2023;86(4):351-353 [FREE Full text] [doi: 10.1097/JCMA.0000000000000900] [Medline: 36791246]
11. Liebrenz M, Schleifer R, Buadze A, Bhugra D, Smith A. Generating scholarly content with ChatGPT: ethical challenges for medical publishing. Lancet Digit Health 2023;5(3):e105-e106 [FREE Full text] [doi: 10.1016/S2589-7500(23)00019-5] [Medline: 36754725]
12. Shen Y, Heacock L, Elias J, Hentel KD, Reig B, Shih G, et al. ChatGPT and other large language models are double-edged swords. Radiology 2023;307(2):e230163 [FREE Full text] [doi: 10.1148/radiol.230163] [Medline: 36700838]
13. Lubowitz JH. ChatGPT, an artificial intelligence chatbot, is impacting medical literature. Arthroscopy 2023;39(5):1121-1122 [FREE Full text] [doi: 10.1016/j.arthro.2023.01.015] [Medline: 36797148]
14. Jianning L, Amin D, Jens K, Jan E. ChatGPT in healthcare: a taxonomy and systematic review. medRxiv Preprint posted online on March 30, 2023. [doi: 10.1101/2023.03.30.23287899]
15. Choudhury A, Shamszare H. Investigating the Impact of User Trust on the Adoption and Use of ChatGPT: Survey Analysis. J Med Internet Res 2023 Jun 14;25:e47184 [FREE Full text] [doi: 10.2196/47184] [Medline: 37314848]
16. Hu PJ, Chau PYK, Sheng ORL, Tam KY. Examining the technology acceptance model using physician acceptance of telemedicine technology. J Manage Inform Syst 2015;16(2):91-112 [FREE Full text] [doi: 10.1080/07421222.1999.11518247]
17. Williams MD, Rana NP, Dwivedi YK. The unified theory of acceptance and use of technology (UTAUT): a literature review. J Enterp Inf Manag 2015;28(3):443-448 [FREE Full text] [doi: 10.1108/jeim-09-2014-0088]
18. McKnight DH, Choudhury V, Kacmar C. The impact of initial consumer trust on intentions to transact with a web site: a trust building model. J Strateg Inf Syst 2002;11(3-4):297-323 [FREE Full text] [doi: 10.1016/S0963-8687(02)00020-3]
19. Hart SG, Staveland LE. Development of NASA-TLX (Task Load Index): results of empirical and theoretical research. Adv Psychol 1988;52:139-183 [FREE Full text] [doi: 10.1016/S0166-4115(08)62386-9]
20. Delone WH, McLean ER. The DeLone and McLean model of information systems success: a ten-year update. J Manag Inf Syst 2014;19(4):9-30 [FREE Full text] [doi: 10.1080/07421222.2003.11045748]
21. Featherman MS, Pavlou PA. Predicting e-services adoption: a perceived risk facets perspective. Int J Hum Comput Stud 2003;59(4):451-474 [FREE Full text] [doi: 10.1016/s1071-5819(03)00111-3]
22. Hair JF, Sarstedt M, Ringle CM, Gudergan SP. Advanced Issues in Partial Least Squares Structural Equation Modeling. 2nd Edition. Thousand Oaks, CA: SAGE Publications; 2017.
23. Eriksson K, Hermansson C, Jonsson S. The performance generating limitations of the relationship-banking model in the digital era—effects of customers' trust, satisfaction, and loyalty on client-level performance. Int J Bank Mark 2020;38(4):889-916 [FREE Full text] [doi: 10.1108/ijbm-08-2019-0282]
24. Al-Ansi A, Olya HGT, Han H. Effect of general risk on trust, satisfaction, and recommendation intention for halal food. Int J Hosp Manag 2019;83:210-219 [FREE Full text] [doi: 10.1016/j.ijhm.2018.10.017]

XSL•FO
RenderX

25. Inegbedion H, Inegbedion E, Peter A, Harry L. Perception of workload balance and employee job satisfaction in work organisations. Heliyon 2020;6(1):e03160 [FREE Full text] [doi: 10.1016/j.heliyon.2020.e03160] [Medline: 31938748]
26. Marikyan D, Papagiannidis S, Rana OF, Ranjan R, Morgan G. "Alexa, let's talk about my productivity": the impact of digital assistants on work productivity. J Bus Res 2022;142:572-584 [FREE Full text] [doi: 10.1016/j.jbusres.2022.01.015]
27. FACT SHEET: Biden- Harris Administration takes new steps to advance responsible artificial intelligence research, development, and deployment. The White House. 2023. URL: https://tinyurl.com/bdfnb97b [accessed 2024-05-07]
28. Tabassi E. Artificial Intelligence Risk Management Framework (AI RMF 1.0). National Institute of Standards and Technology. 2023. URL: https://nvlpubs.nist.gov/nistpubs/ai/NIST.AI.100-1.pdf [accessed 2024-05-07]
29. Sato T, Yamani Y, Liechty M, Chancey ET. Automation trust increases under high-workload multitasking scenarios involving risk. Cogn Tech Work 2019;22(2):399-407 [FREE Full text] [doi: 10.1007/s10111-019-00580-5]
30. Karpinsky ND, Chancey ET, Palmer DB, Yamani Y. Automation trust and attention allocation in multitasking workspace. Appl Ergon 2018;70:194-201 [FREE Full text] [doi: 10.1016/j.apergo.2018.03.008] [Medline: 29866311]
31. Gontar P, Homans H, Rostalski M, Behrend J, Dehais F, Bengler K. Are pilots prepared for a cyber-attack? A human factors approach to the experimental evaluation of pilots' behavior. J Air Transp Manag 2018;69:26-37 [FREE Full text] [doi: 10.1016/j.jairtraman.2018.01.004]
32. Tentama F, Rahmawati PA, Muhopilah P. The effect and implications of work stress and workload on job satisfaction. Int J Sci Technol Res 2019;8(11):2498-2502 [FREE Full text]
33. Kim C, Mirusmonov M, Lee I. An empirical examination of factors influencing the intention to use mobile payment. Comput Human Behav 2010;26(3):310-322 [FREE Full text] [doi: 10.1016/j.chb.2009.10.013]
34. Holland P, Tham TL, Sheehan C, Cooper B. The impact of perceived workload on nurse satisfaction with work-life balance and intention to leave the occupation. Appl Nurs Res 2019;49:70-76 [FREE Full text] [doi: 10.1016/j.apnr.2019.06.001] [Medline: 31375315]
35. Nguyen TM, Malik A. A two-wave cross-lagged study on AI service quality: the moderating effects of the job level and job role. Br J Manag 2021;33(3):1221-1237 [FREE Full text] [doi: 10.1111/1467-8551.12540]
36. Kumar A, Adlakaha A, Mukherjee K. The effect of perceived security and grievance redressal on continuance intention to use M-wallets in a developing country. Int J Bank Mark 2018;36(7):1170-1189 [FREE Full text] [doi: 10.1108/ijbm-04-2017-0077]
37. Chen X, Li S. Understanding continuance intention of mobile payment services: an empirical study. J Comput Inf Syst 2016;57(4):287-298 [FREE Full text] [doi: 10.1080/08874417.2016.1180649]
38. Fang Y, Qureshi I, Sun H, McCole P, Ramsey E, Lim KH. Trust, satisfaction, and online repurchase intention. MIS Q 2014;38(2):407-428 [FREE Full text]
39. Gu Z, Wei J, Xu F. An empirical study on factors influencing consumers' initial trust in wearable commerce. J Comput Inf Syst 2015;56(1):79-85. [doi: 10.1080/08874417.2015.11645804]
40. Oliveira T, Faria M, Thomas MA, Popovič A. Extending the understanding of mobile banking adoption: when UTAUT meets TTF and ITM. Int J Inform Manage 2014;34(5):689-703 [FREE Full text] [doi: 10.1016/j.ijinfomgt.2014.06.004]
41. Mostafa RB, Kasamani T. Antecedents and consequences of chatbot initial trust. Eur J Mark 2021;56(6):1748-1771 [FREE Full text] [doi: 10.1108/ejm-02-2020-0084]
42. Wang X, Lin X, Shao B. Artificial intelligence changes the way we work: a close look at innovating with chatbots. J Assoc Inf Sci Technol 2022;74(3):339-353 [FREE Full text] [doi: 10.1002/asi.24621]
43. Hsiao KL, Chen CC. What drives continuance intention to use a food-ordering chatbot? An examination of trust and satisfaction. Libr Hi Tech 2021;40(4):929-946 [FREE Full text] [doi: 10.1108/lht-08-2021-0274]
44. Pesonen JA. 'Are You OK?' Students' trust in a chatbot providing support opportunities. : Springer; 2021 Presented at: Learning and Collaboration Technologies: Games and Virtual Environments for Learning: 8th International Conference, LCT 2021, Held as Part of the 23rd HCI International Conference, HCII 2021; July 24-29, 2021; Virtual Event. [doi: 10.1007/978-3-030-77943-6_13]
45. Dwivedi YK, Kshetri N, Hughes L, Slade EL, Jeyaraj A, Kar AK, et al. Opinion paper: "So what if ChatGPT wrote it?" Multidisciplinary perspectives on opportunities, challenges and implications of generative conversational AI for research, practice and policy. Int J Inform Manage 2023;71:102642 [FREE Full text] [doi: 10.1016/j.ijinfomgt.2023.102642]
46. Silva SC, De Cicco R, Vlačić B, Elmashhara MG. Using chatbots in e-retailing—how to mitigate perceived risk and enhance the flow experience. Int J Retail Distrib Manag 2022;51(3):285-305 [FREE Full text] [doi: 10.1108/ijrdm-05-2022-0163]
47. Nordheim CB, Følstad A, Bjørkli CA. An initial model of trust in chatbots for customer service? Findings from a questionnaire study. Interact Comput 2019;31(3):317-335 [FREE Full text] [doi: 10.1093/iwc/iwz022]

Abbreviations

AI: artificial intelligence
AI RMF: artificial intelligence risk management framework
PLS-SEM: partial least squares structural equation modeling
RMSEA: root mean square error of approximation

XSL•FO
RenderX

TAM: technology acceptance model
UTAUT: unified theory of acceptance and use of technology

Edited by A Kushniruk, E Borycki; submitted 11.12.23; peer-reviewed by P Radanliev, G Farid; comments to author 17.01.24; revised version received 25.03.24; accepted 07.04.24; published 27.05.24.

Please cite as:
Choudhury A, Shamszare H
The Impact of Performance Expectancy, Workload, Risk, and Satisfaction on Trust in ChatGPT: Cross-Sectional Survey Analysis
JMIR Hum Factors 2024;11:e55399
URL: https://humanfactors.jmir.org/2024/1/e55399
doi:10.2196/55399
PMID:38801658

©Avishek Choudhury, Hamid Shamszare. Originally published in JMIR Human Factors (https://humanfactors.jmir.org), 27.05.2024. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on https://humanfactors.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

The Role of Coherent Robot Behavior and Embodiment in Emotion Perception and Recognition During Human-Robot Interaction: Experimental Study

Laura Fiorini[1,2], PhD; Grazia D'Onofrio[3], PhD; Alessandra Sorrentino[1], PhD; Federica Gabriella Cornacchia Loizzo[2], MSc; Sergio Russo[4], MSc; Filomena Ciccone[3], PsyD; Francesco Giuliani[4], MSc; Daniele Sancarlo[5], MD; Filippo Cavallo[1,2], PhD

[1]Department of Industrial Engineering, University of Florence, Firenze, Italy

[2]The BioRobotics Institute, Scuola Superiore Sant'Anna, Pontedera (Pisa), Italy

[3]Clinical Psychology Service, Health Department, Fondazione IRCCS Casa Sollievo della Sofferenza, San Giovanni Rotondo (Foggia), Italy

[4]Innovation & Research Unit, Fondazione IRCCS Casa Sollievo della Sofferenza, San Giovanni Rotondo (Foggia), Italy

[5]Complex Unit of Geriatrics, Department of Medical Sciences, Fondazione IRCCS Casa Sollievo della Sofferenza, San Giovanni Rotondo (Foggia), Italy

Corresponding Author:
Laura Fiorini, PhD
Department of Industrial Engineering
University of Florence
Via Santa Marta 3
Firenze, 50139
Italy
Phone: 39 0552758663
Email: laura.fiorini@unifi.it

Abstract

Background: Social robots are becoming increasingly important as companions in our daily lives. Consequently, humans expect to interact with them using the same mental models applied to human-human interactions, including the use of cospeech gestures. Research efforts have been devoted to understanding users' needs and developing robot's behavioral models that can perceive the user state and properly plan a reaction. Despite the efforts made, some challenges regarding the effect of robot embodiment and behavior in the perception of emotions remain open.

Objective: The aim of this study is dual. First, it aims to assess the role of the robot's cospeech gestures and embodiment in the user's perceived emotions in terms of valence (stimulus pleasantness), arousal (intensity of evoked emotion), and dominance (degree of control exerted by the stimulus). Second, it aims to evaluate the robot's accuracy in identifying positive, negative, and neutral emotions displayed by interacting humans using 3 supervised machine learning algorithms: support vector machine, random forest, and K-nearest neighbor.

Methods: Pepper robot was used to elicit the 3 emotions in humans using a set of 60 images retrieved from a standardized database. In particular, 2 experimental conditions for emotion elicitation were performed with Pepper robot: with a static behavior or with a robot that expresses coherent (COH) cospeech behavior. Furthermore, to evaluate the role of the robot embodiment, the third elicitation was performed by asking the participant to interact with a PC, where a graphical interface showed the same images. Each participant was requested to undergo only 1 of the 3 experimental conditions.

Results: A total of 60 participants were recruited for this study, 20 for each experimental condition for a total of 3600 interactions. The results showed significant differences ($P<.05$) in valence, arousal, and dominance when stimulated with the Pepper robot behaving COH with respect to the PC condition, thus underlying the importance of the robot's nonverbal communication and embodiment. A higher valence score was obtained for the elicitation of the robot (COH and robot with static behavior) with respect to the PC. For emotion recognition, the K-nearest neighbor classifiers achieved the best accuracy results. In particular, the COH modality achieved the highest level of accuracy (0.97) when compared with the static behavior and PC elicitations (0.88 and 0.94, respectively).

XSL•FO
RenderX

Conclusions: The results suggest that the use of multimodal communication channels, such as cospeech and visual channels, as in the COH modality, may improve the recognition accuracy of the user's emotional state and can reinforce the perceived emotion. Future studies should investigate the effect of age, culture, and cognitive profile on the emotion perception and recognition going beyond the limitation of this work.

(JMIR Hum Factors 2024;11:e45494) doi:10.2196/45494

KEYWORDS

social robot; emotion recognition; human emotion perception; human-robot interaction; robot cospeech gestures evaluation

Introduction

Background

During the last decade, there has been increasing interest in research on socially assistive robotics aimed at realizing intelligent robotic solutions for health care and social assistance. We experience an evolution of social robot applications; indeed, they moved from the role of *concierge* and *helper* [1] toward the role of *companion* and *therapist* [2,3]. Social robots have the potential to contribute to the greater good of society; indeed, it has been demonstrated that they can support everyday life as companions and the health care system from *logistics* to *assistance and rehabilitation*. Thinking to include social robots in the care chain, they can be used to reduce stress, anxiety, and pain in children [4]; they can be integrated into conventional behavioral and cognitive therapies for both children and adults who struggle with social anxiety [5]; or they can be used to promote mental health [6]. A review by Hung et al [7] showed evidence that Paro robots can reduce negative emotions in patients, promoting a positive mood and improving social engagement. Rossi et al [8] demonstrated that social robots are effective in decreasing stress in children accessing the emergency room. As the complexity of the robot task increases, social robots are required to perform more complex perceptual, cognitive, and interactive functionalities. This is the case in long-term interactions in which robots and users should establish meaningful communication, emotional awareness, and reliable engagement.

In this context, the human-robot interaction (HRI) field has become crucial, and it is now compelling to better understand how humans perceive, interact with, or accept these machines in social and real contexts. Researchers are also debating on defining the factors that can influence the perceived social capabilities and intelligence of a robot [9,10]. De Graaf et al [11] highlighted the significance of the robot's social capability, emphasizing the importance of 2-way interaction where a robot is expected to respond to humans in a social manner. In addition, De Graaf et al [11] underlined that a social robot should *also display thoughts and feelings* and should be *socially aware of the environment*, among other issues. When a robot failed to perform this 2-way interaction, people were disappointed and experienced a sense of dissonance. In other words, when interacting with a social robot, especially a humanoid robot, we expect to use the same mental structure and social rules that guide us in human-human communication, expecting empathetic interaction because they are perceived as social actors [12].

From a roboticist or engineering point of view, these concepts are translated into the design and development of behavioral models that can guarantee an efficient and reliable 2-way interaction [13,14]; they should *perceive* and *show* emotions (and social norms) and thus be understood by humans with whom they are interacting. The key challenge in this field is to provide robots with cognitive and affective capabilities, developing architectures that allow them to establish empathetic relationships with users, which can foster long-term and meaningful interactions. From an implementation perspective, the design and the deployment of a socially capable social robot comprises 2 essential parts. The first is devoted to designing and implementing a consistent and congruent emotional behavioral architecture that makes the robot react or act to the environment (ie, *display thoughts and feelings*). The capabilities of a user to understand the emotions displayed by a robot have been explored in different settings [15,16]. Examples of actions can include the expression of congruent cues such as facial expressions [17], changes in the color of the eyes, movement of the upper limbs [16,18], or smart navigation strategies [19]. In contrast, the other part is more focused on the robot's perception of the user's emotional response to these behaviors [20], with special attention to contextualizing its action and reaction according to the living contexts and habits or preferences of the person with whom it is interacting (ie, *being socially aware of the environment*) [21].

Related Work on Emotion and Social Robots

The ability of a robot to perceive the nonverbal cues of the user, which convey user emotion and intent, plays a key role in the development of social robots capable of performing meaningful interactions [22,23]. In this sense, humans' gaze, body posture, cospeech gestures, and facial expressions play a leading role in defining the context of the interaction, helping the robot to correctly classify the experience, and associating it with informative content [21]. The development of such abilities, for a researcher in the field of robotics, translates into the use of multimodal sensor modality and the implementation of several complex algorithms to endow robots with different cognitive and social capabilities. The visual modality is the most commonly used [24] because it can detect nonverbal behaviors that are representative of the emotional state of users without requiring them to wear any external sensor. Alternatively, wearable sensors [25] can be used, also using a multimodal approach, to overcome the problems related to occlusion and low light. Other algorithms or modules were implemented to perform multiperson tracking [26], speech recognition [27,28], and automatic engagement detection [29]. A recent review paper [24] provides a deep insight into the most used methods and approaches.

XSL•FO
RenderX

For the *showemotion* part, robots must exploit several channels (ie, auditory, visual, cospeech, and gestures) and mechanisms (eg, body posture, facial expressions, vocal prosody, touch, and gaze) to communicate their "internal emotional status" and intentions authentically and clearly [30]. Thus, the capabilities of a user to understand the emotions displayed by a robot have been explored in several settings [31]. Over the last few years, several attempts have been made using both video-simulated robots and real robots. Guo et al [20] showed participants 5 different emotions using the humanoid robot called Alpha2, and they were asked to rate the perceived emotion using the Self-Assessment Manikin questionnaire (SAM; only valence and arousal dimensions) [32]. In contrast, Barchard et al [33] conducted a web-based study to evaluate the perception of a robot's social intelligence by showing videos of robot interactions. However, the embodiment and the appearance of social robots play important roles in the perception of the robot; therefore, video-based elicitation could introduce some bias in the analysis of perceived emotion. This is why other research has relied on investigating the emotion perceived during a real HRI. This is the case of Bagheri et al [34], who asked participants to watch 6 performances of America's Got Talent Show on Pepper's tablet that are expected to evoke the 6 basic emotions. Rossi et al [35] and Staffa et al [36] relied on movie trailers to evoke emotions. However, they used nonstandard videos, making it challenging to identify the target emotion in a recognized and standardized manner, as the elicited emotion through the video clips is not known a priori, and consequently, it is difficult to define the role of the robot (and its embodiment) in the elicitation process.

Research groups have recently begun to study the effects of multimodal channels on communication. Studies conducted with embodied conversational agents showed that incongruent emotional stimuli (eg, auditory and visual stimuli) can result in adverse consequences on user rating; conversely, congruent stimuli can facilitate the recognition of emotions [37]. Other researchers have also studied the role of nonverbal behavioral cues while interacting with robots. Movie clips showing coherent and incoherent robot behaviors are often used to elicit emotional responses from users with respect to those induced by movie clips [15,16,18,35]. For instance, Rossi et al [16] investigated how an incoherent nonverbal robot's behavior with respect to the presented emotion can produce a type of humorous effect. Tsiourti et al [18] investigated how contextual incongruence (ie, a robot's reaction conflicts with the socioemotional context) can confuse the observers, decreasing the accuracy of the perceived emotion. Nevertheless, such a cospeech robot's behavior was used in addition to a nonstandard method of emotion elicitation, as previously remarked; thus, it is not easy to understand the role of the robot's behavior with respect to the emotional context. Therefore, it is important to understand how the robot's nonverbal behavior might shape the human perception of the showed emotion elicited through standard emotionally labeled visual data sets and, at the same time, observe the robot's emotion recognition accuracy rate. Although previous studies have shown a correlation between the robot's nonverbal action and perceived emotion, there is a lack of use of standard elicitation modalities.

Therefore, in this work, we present the results of 3 experimental sessions to observe the performance of the robot in recognizing users' emotions as well as to investigate the difference (if any) in eliciting emotions in humans when using a social robot (with or without coherent behavior) rather than a PC. We plan to use a standard data set of pictures, namely, the International Affective Picture System (IAPS) [38], to elicit emotions in users. Particularly, the robot will use a multimodal behavior (ie, head movements, vocal reinforcement, and body gestures) to interact with the participants while showing the graphical emotions by establishing social binding, whereas the PC will provide emotion elicitation only through a graphical interface. The 2 graphical interfaces have been designed to provide the same information to the user but using different communication channels. In this context, the aim of this work is dual. First, it aims to investigate the increase in the user's emotional perception during the interaction with a robot with respect to a PC (Figure 1, blue arrow). In particular, this work investigates the role of the robot's coherent nonverbal behavior in emotion perception by consequently assessing the impact of robot embodiment and, eventually, its coherent behavior. Robot nonverbal cues are manipulated with respect to a mapping between the main associated emotion and cospeech gestures that can be generated on the robot. At the end of each interaction, the participants were asked to self-assess their perceived emotions. In this study, we used the emotion classification proposed by Russel et al [39], which relies on 3 variables, namely, valence, arousal, and dominance. Valence describes the degree to which a stimulus causes a positive or negative emotion, arousal refers to the intensity or level of energy invested in the emotion, and dominance reflects the extent of perceived control over the emotional response when facing the stimulus. The collected answers were analyzed to answer the following research questions (RQs):

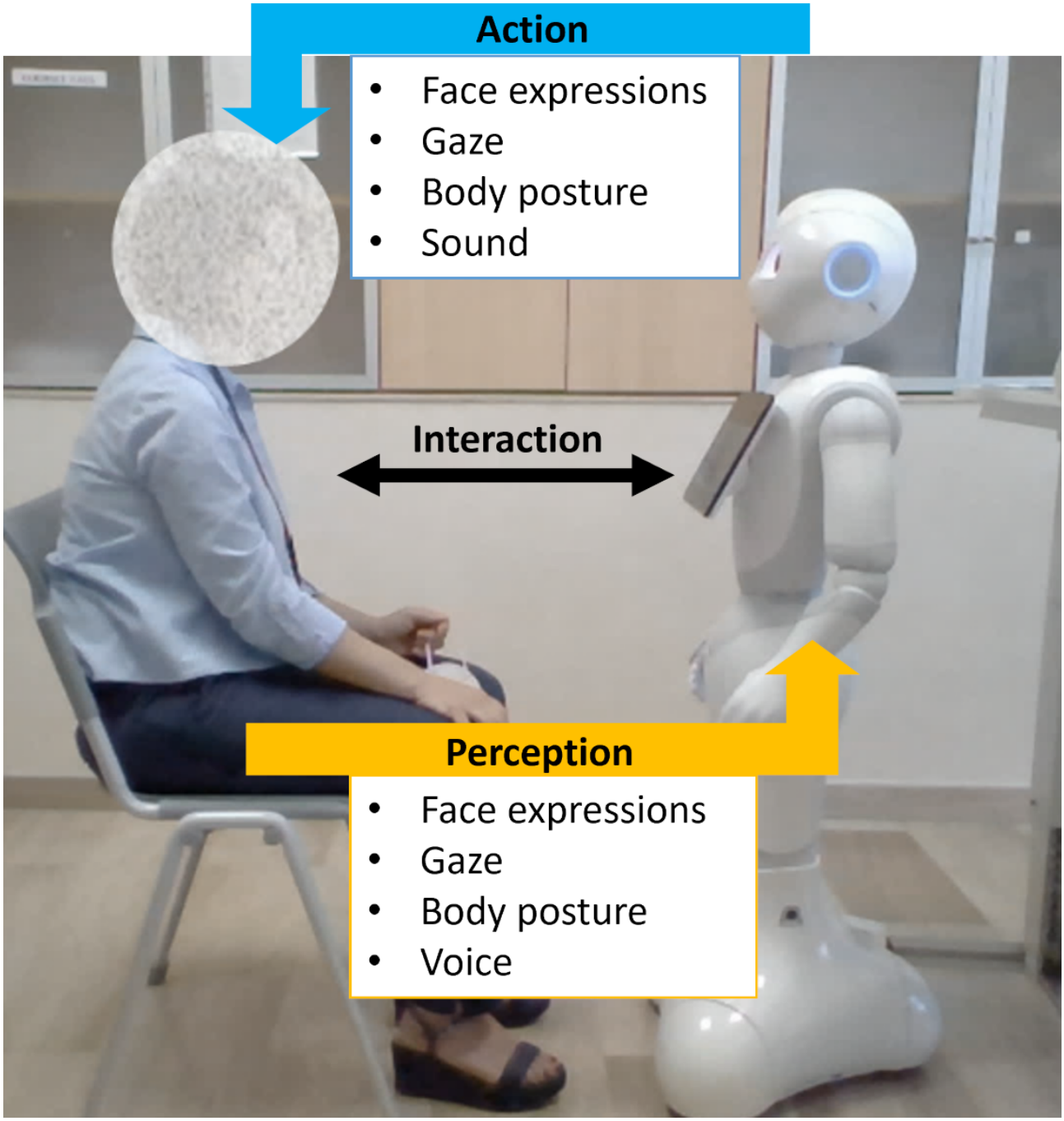
1. RQ1: Emotion elicited through a humanoid robot interacting with coherent emotional behavior is rated higher than emotions elicited by a web application in terms of emotional valence, arousal, and dominance.
2. RQ2: There are significant differences in terms of emotional valence, arousal, and dominance between a robot showing coherent behavior rather than a robot that it is not moving at all (static condition).
3. RQ3: The embodiment of the humanoid robot will not affect the emotion perception compared with the web application.

Second, this study aims to assess the accuracy of the robot in recognizing the elicited emotion in the participants (Figure 1, yellow arrow). The ability to infer and interpret emotions plays a key role in establishing intuitive and engaging HRIs. On the one hand, a robot endowed with emotion recognition skills can adapt its behavior based on the detected user emotion [22]. On the other hand, a robot expressing recognizable emotions positively influences the evaluation of its capabilities [40]. In particular, features related to facial expressions were extracted, preprocessed, and analyzed with 3 supervised machine learning techniques to verify the following RQ:

1. RQ4—There is no difference in the robot emotion recognition accuracy despite the elicitation modalities (robot or web application).

XSL•FO
RenderX

Figure 1. Two-way interaction proposed in this study. To improve the human-robot interaction, the robot should perceive the user's behavior (yellow arrow) and plan appropriate action (blue arrow).



In our previous studies [41,42], we evaluated the perceived acceptance and the recognition rate of having a robot that acts coherently and incoherently despite the standard emotion showed with respect to the standard elicitation modality. In contrast, in this study, we focus only on coherent behavior by comparing it with a standard web application that runs on a PC. In addition, instead of focusing on evaluating how the robot's acceptance is modulated according to the elicitation modality, we focused on the perceived emotion evoked.

Methods

Instrumentation

The instrumentation is composed of the following elements: (1) a Pepper robot (Aldebaran, United Robotics Group) or a PC, (2) the RoboMate (Behaviour Labs) interface for cospeech gestures, (3) a custom interface that contains pictures from the IAPS for eliciting emotion, and (4) an external camera placed on Pepper to record the participants' emotions during the interaction. Pepper is a humanoid robot that is widely used for experimentation in socially assistive robotics. It is 120 cm tall, weighs 28 kg, and has 20 df, including 1 head, 2 arms, and 1 wheeled base. In addition, it has a tablet on the front. Robot

XSL•FO
RenderX

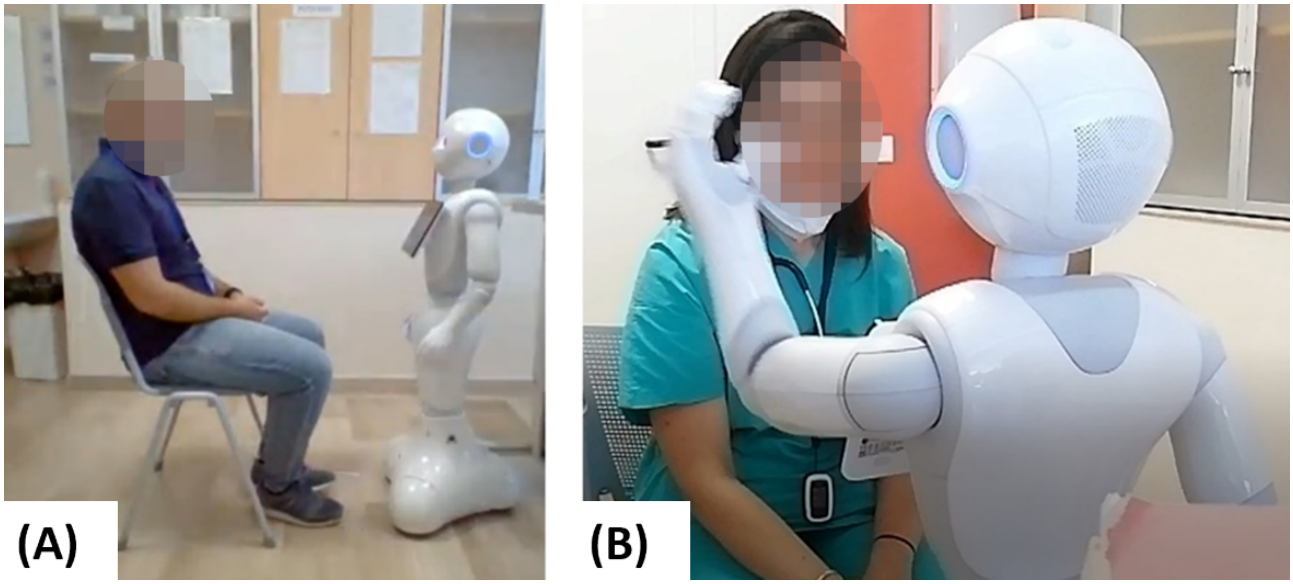
coherent behavior was managed through the RoboMate interface [43] to animate Pepper, when necessary, selecting among the behaviors classified as "positive social stimulus" or "negative social stimulus." The selected stimulus was modeled by a psychologist using 3 modalities: body gestures (upper limb and head), gaze, and sound. IAPS is a database of images devoted to eliciting standardized emotions [44]. It was developed by the Center for Emotion and Attention at the University of Florida. This database is commonly used in psychological studies on emotions and attention. Each image in the data set is labeled with the corresponding emotion, thus enabling researchers to properly select the stimulus. In this study, 60 images were selected from the team of psychologists of the hospital "Casa Sollievo della Sofferenza." According to the IAPS valence dimension, 21 of the selected images were rated as positive, 19 as negative, and 20 as neutral. A customized web-based interface was developed to standardize the emotional stimulation when using 2 different communication channels (a robot and a PC).

Experimental Setup

A psychologist welcomed the participant, briefly explaining the experimental setup, including how to use the evaluation tool. It is important to emphasize that the participant was not aware of the real objective of the experimentation, thus avoiding interference with the experience. To properly investigate the RQs, each participant underwent 1 of the following elicitation modalities.

1. Static (STA) behavior: Pepper robot has its arms along the body in a neutral position (Figure 2A). Pepper's face was looking at the participant but without any animacy. Pepper displayed IAPS images on its tablet through the customized web application.
2. Coherent (COH) behavior: Similar to the STA condition, the IAPS images were shown on Pepper's tablet. Using the RoboMate application, the psychologist assigned a coherent behavior to Pepper with the shown images. In particular, the psychologist can choose and combine 3 modalities for elicited emotions: body gesture (upper limb and head), gaze, and sound, which are available on the RoboMate application (Figure 2B). For example, in the case of positive emotion, Pepper's gestures were chosen to look friendly; it should look to the user direction, and the voice gave positive reinforcements.
3. PC: For this experimental condition, we used a PC instead of the Pepper robot. Participants were asked to evaluate the images shown on a PC through the customized web application.

Figure 2. Experimental setup. The participants were interacting with Pepper robot during the experimentation. (A) Participants were asked to sit in front of the robot and watch the images on its tablet. (B) If the participant belonged to the coherent elicitation modality group, the Pepper robot would move its arms, eyes, and head.



The participant was asked to sit in front of the technology (ie, Pepper robot or PC). If the user interacts with Pepper, Pepper is placed 0.5 to 0.6 m far from the user (ie, personal distances [45]); in the case of interaction with the PC, the user is requested to sit and interact with the computer as he or she will commonly do.

Each stimulus was shown for 7 seconds, and at the end, the participant was asked to fill out the SAM [32], as adapted in the study by Gatti et al [46] directly on the robot or on the computer after each picture. SAM is an emotion assessment tool that uses graphic scales, depicting cartoon characters expressing 3 emotional elements (valence, arousal, and dominance). Each participant was asked to rate the domains by selecting an image that corresponded to a score between 1 and 9. A picture of the interface is presented in Multimedia Appendix 1.

At the end of the experimental session, each participant completed 60 SAM questionnaires. The psychologist was present during the test, and she or he was ready to intervene in case of necessity. All the tests were performed at the "Casa Sollievo della Sofferenza" research hospital.

Ethical Considerations

The approval of the study for experiments using human participants was obtained from the local Ethics Committee on Human Experimentation (register code 3038/01DG). All

XSL•FO
RenderX

participants signed an informed consent form before participating in this study, and pictured participants provided written informed consent to allow their image to be published. The data were pseudoanonymized and stored on a GDPR-compliant server.

Participants

Participants were recruited from July 2020 to February 2021 from employees and staff of the "Casa Sollievo della Sofferenza" research hospital located in Apulia (San Giovanni Rotondo, Foggia) using convenience sampling. Participants were excluded if they had a hearing or visual impairment. Recruited participants were then randomly assigned to undergo 1 of the 3 experimental conditions (ie, STA, COH, and PC). Sociodemographic information (age, education, and sex) was collected to verify the similarities between the groups.

Data Analysis

Overview

Owing to the sample size of each cohort, the nonparametric statistic was used, particularly the Kruskal-Wallis test and chi-square test, to investigate significant differences between participants' groups in terms of age, sex, and educational level. The significance level was set at P=.05. The following paragraphs describe the analysis performed on the SAM questionnaires and the data collected from camera sensors.

Emotion Perception Analysis

A total of 60 SAM questionnaires were collected for each participant. The average values of the valence, arousal, and dominance domains were computed for each selected image of each group of elicitation modality (ie, STA, COH, and PC). Differences were analyzed with the Kruskal-Wallis test (P<.05) and post hoc evaluated with the Mann-Whitney U test (with Bonferroni correction) used to identify between which pair of elicitation modes the difference has occurred.

Emotion Recognition Analysis

Data from the camera were processed and examined offline. The recordings were initially analyzed [47] to ensure that only the frames featuring the face of the person performing the test were included in the study. Then the recordings were segmented, providing short videos that corresponded to the user's reaction to each image proposed, totaling 60 videos per user. The OpenFace toolkit [48] was used to extract 150 features related to gaze and facial expression from each video as well as the quality (ie, confidence) of the extracted features. The data were filtered according to the confidence score (frames with a confidence score <0.90 were discarded). The data were then labeled based on the IAPS-defined emotions (ie, positive, negative, and neutral). Data were normalized and selected. Only features with a correlation coefficient of <0.85 were picked from the initial data set, avoiding those with a high correlation coefficient (which may represent redundant information). The data of the merged data set were then separated into sub–data sets (one for each participant), and emotion classification was performed using the selected features. In this study, we rely on state-of-the-art methods used for emotion recognition [24] to facilitate a comparison with other works. The 3 supervised classifiers used are support vector machine (SVM), random forest (RF), and K-nearest neighbor (KNN). To classify the data by participant, a 10-fold cross-validation procedure was applied, and the outputs were organized in a confusion matrix. The classification performance was assessed in terms of accuracy, precision, recall, and F-measure [49]. The calculations were computed using MATLAB 2020a. More details on emotion recognition analysis are available in Multimedia Appendix 2 [24,47-49].

Results

Description of the Participant Cohort

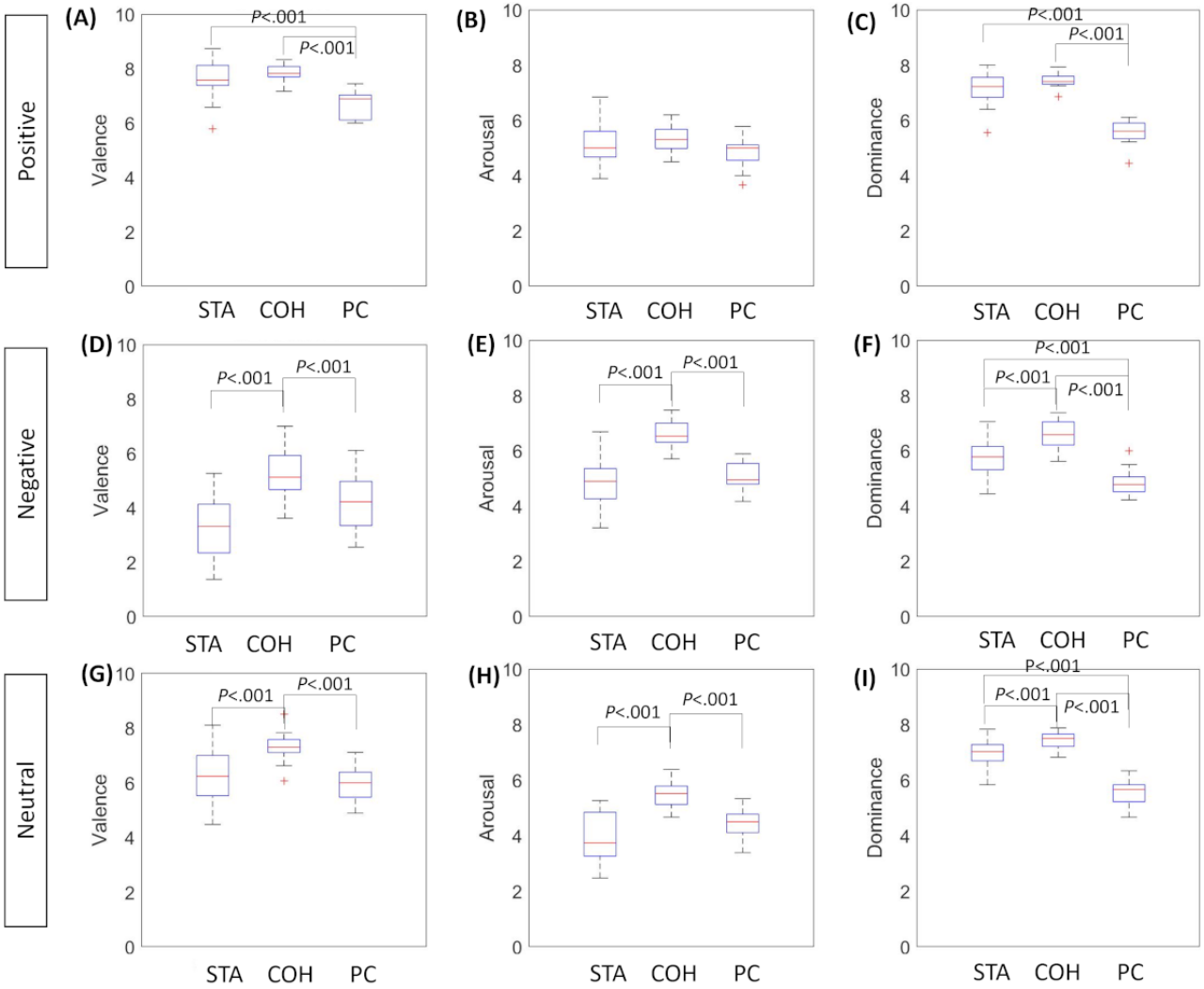
A total of 60 participants were involved in this study, 20 for each modality, resulting in 3600 interactions with technologies. In total, 3 participants were excluded from the analysis of perceived emotion because not all SAM evaluations were correctly saved after each elicitation. In case of missing SAM values, these ratings were removed from the analysis of average values. Finally, 57 participants were included in these subgroups of analyses linked to RQ1, RQ2, and RQ3. Regarding the recognition of emotion using machine learning techniques (linked to RQ4), a total of 53 participants were included in the analysis. A total of 7 participants were excluded because of technical problems related to the quality of the recorded images. The statistical tests did not indicate any difference between the 3 participant cohorts regarding age, sex, and educational level. The participant demographics and educational analyses are reported in Multimedia Appendix 3.

Participants' Perceived Emotion Results

The results underline significant differences (P<.001) in the perceived emotions according to the different elicitation modalities, except for the arousal elicited with the positive images (Figure 3). The median and IQR values are fully reported in Multimedia Appendix 4. As for valence, the robot with coherent behavior elicited significant differences (P<.001) and higher values in terms of valence, arousal, and dominance domains compared with the other 2 modalities for negative and neutral emotions. In terms of negative valence, the participants perceived fewer negative emotions with the coherent robot than with the other 2 modalities. For positive valence, elicitation with the web application is significantly different from that with the robot (P<.001).

XSL•FO
RenderX

Figure 3. Self-Assessment Manikin Questionnaire results for the 3 elicitation modalities. Boxplot matrix (A), (B), and (C) denote valence, arousal, and dominance for the positive elicitation, respectively; (D), (E), and (F) denote valence, arousal, and dominance for the negative elicitation, respectively; (G), (H), and (I) denote valence, arousal, and dominance for the neutral elicitation, respectively; asterisks on boxplot remark the significant differences evaluated with the Mann-Whitney post hoc test corrected with Bonferroni. COH: coherent; STA: static.



Regarding arousal, the coherent robot was rated higher than the other 2 modalities, but there were significant differences ($P<.001$) only for negative and neutral emotions, whereas for positive arousal, the results, depicted in Figure 3, highlight only a trend. All the P values are reported in Multimedia Appendix 4.

The participants stimulated using the robot rated significantly higher dominance across all 3 emotions rather than the cohort that used the PC in the test. As for positive elicitation, we found significant differences ($P<.001$) between the cohort stimulated with the PC and those stimulated with the robot (ie, static behavior and coherent behavior). Indeed, the participants rated the emotions (in terms of valence and arousal) elicited by the robot more than the ones elicited using the PC. All P values are reported in Multimedia Appendix 4.

Robot's Emotion Recognition Results

Because of technical issues 1848 frames pertaining to the PC modality were removed from the analysis during the preprocessing. At the end, the total number of samples included in this study was 296,677 for the STA modality, 228,170 for the COH modality, and 103,758 for the PC modality. The number of columns in each data set corresponded to the number of features selected using the correlation analysis method. The following features were selected (Figure 4):

1. The x-, y-, and z-coordinates of the eye gaze direction vector for eye 0 (3 features).
2. The z-coordinate of the eye gaze direction vector for eye 1 (1 feature).
3. The x- and y-coordinates of the location of the landmark 8 (the leftmost in the image) of the eye 0 (2 features).

The 53 data sets were fed into 3 classifiers (SVM, RF, and KNN) [24]. The data sets were uniformly distributed across the 3 groups, as presented in Table 1.

XSL•FO
RenderX

Figure 4. Selected features. (A) Face and (B) eye landmarks extracted with OpenFace software. The landmark 8 in panel B was chosen after the feature selection.

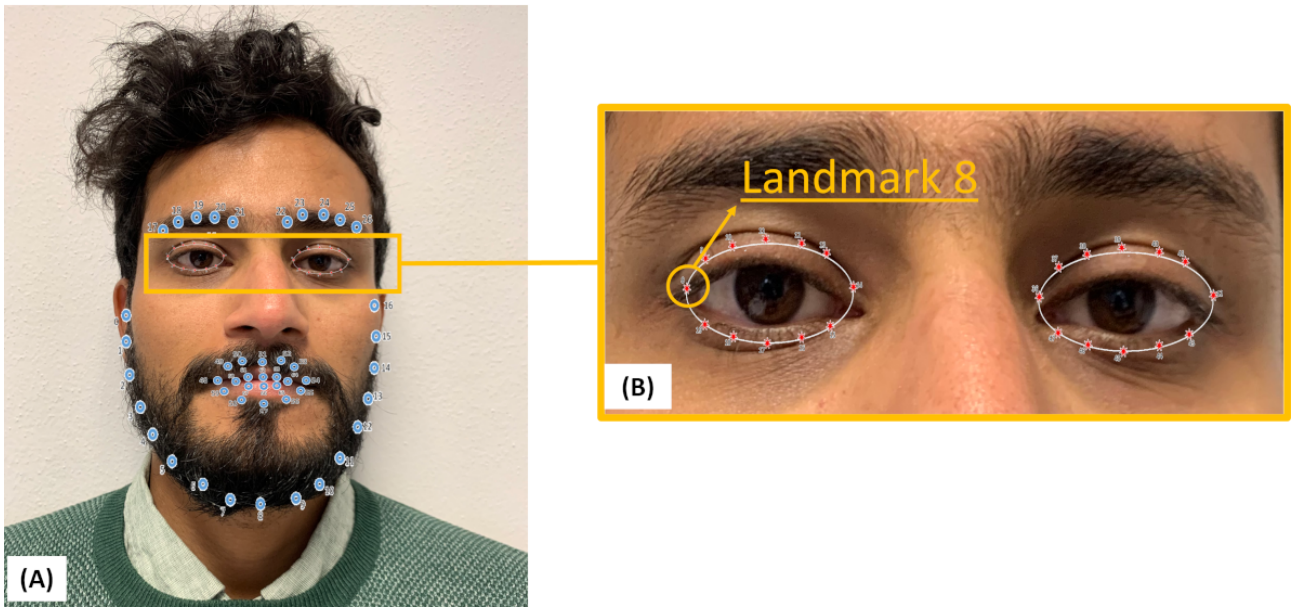


Table 1. Distribution of data set instances.

Group	Positive, n (%)	Negative, n (%)	Neutral, n (%)
Static (n=296,677)	103,992 (35.03)	94,257 (31.77)	98,499 (33.2)
Coherent (n=228,170)	70,710 (30.99)	74,383 (32.6)	83,077 (36.41)
PC (n=103,758)	35,195 (33.92)	32,072 (30.91)	36,492 (35.17)

Accuracy, precision, F-measure, and recall were calculated as the mean values from the participants in the same experimental cohort. According to the findings, the KNN classifier offers the best classification results, with an accuracy of up to 0.88 for STA behavior, 0.97 for COH, and 0.94 for PC. The SVM classifiers, in contrast, had the lowest results (accuracy of up to 0.57, 0.67, and 0.68 for STA, COH, and PC, respectively); hence, they were excluded from further research. Compared with the RF classifier, the KNN classifier has the best F-measure (>0.88).

Table 2 presents the complete results for the KNN and RF classifiers, including the accuracy, F-measure, precision, and recall for each group. According to the overall trend, the COH modality achieves a high level of accuracy when compared with the STA and PC elicitations. In terms of the other indicators, the COH was better with the KNN classifier and slightly worse with the RF classifier when it came to elicitation with the PC.

Table 2. Performance of K-nearest neighbor (KNN) and random forest (RF) classifiersa.

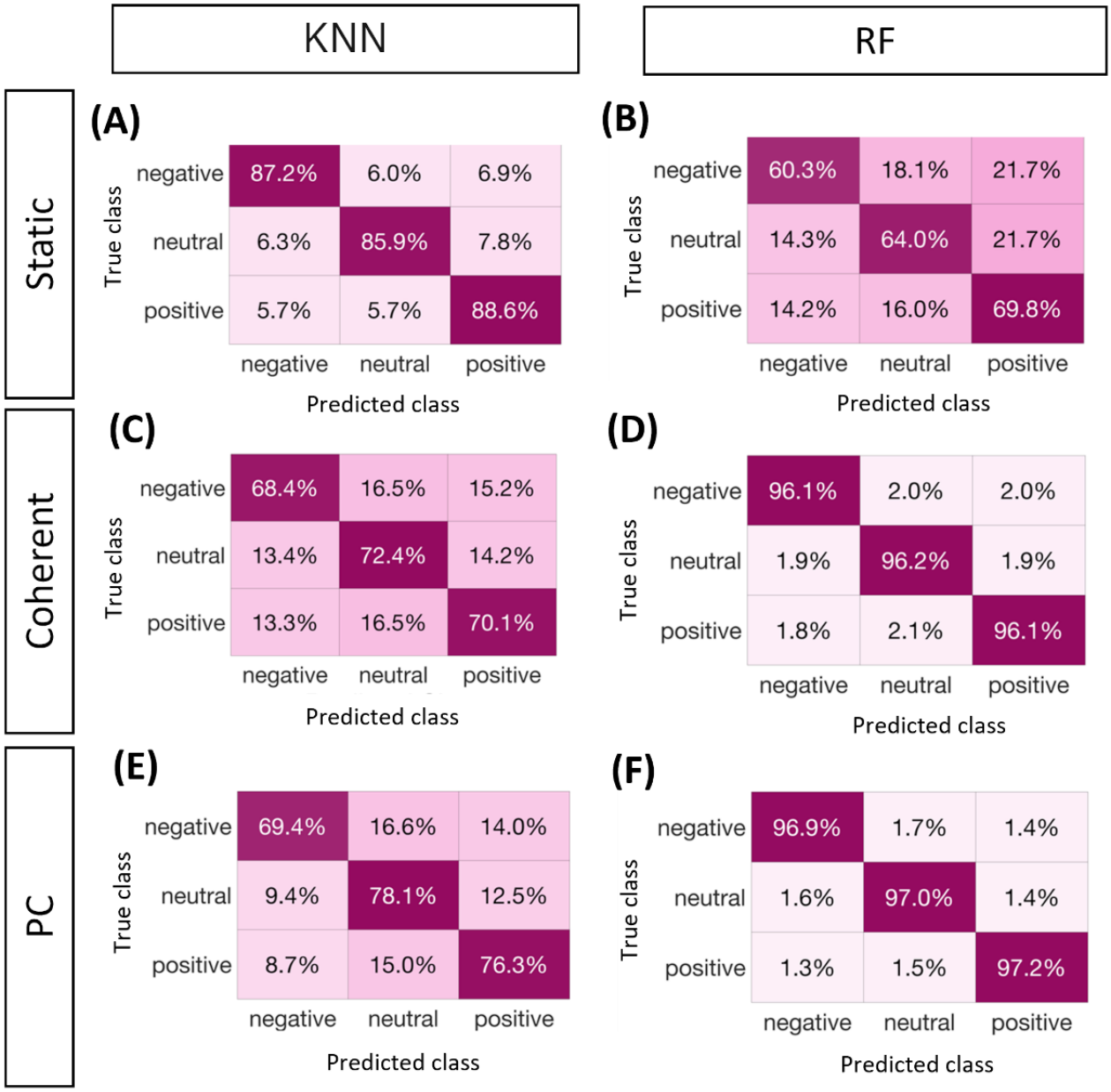
Group	Accuracy		Precision		F-measure		Recall	
	KNN	RF	KNN	RF	KNN	RF	KNN	RF
Static	0.88	0.65	0.88	0.65	0.88	0.65	0.88	0.65
Coherent	0.97	0.73	0.96	0.72	0.96	0.72	0.96	0.72
PC	0.94	0.74	0.94	0.74	0.94	0.74	0.94	0.74

[a]Mean values are used to calculate the results.

Confusion matrices (Figure 5) for the 3 elicitation modalities were generated to investigate the performance of the classifiers in recognizing the 3 selected emotions. The positive emotion was often better identified, whereas the negative emotion was the least recognized. When the user is stimulated with the robot with coherent modality and the PC, the RF classifier performs better than the KNN classifiers in distinguishing emotions. The KNN classifier appeared to perform better in the static modality than in the other 2.

XSL•FO
RenderX

Figure 5. Confusion matrices for K-nearest neighbor (KNN) and random forest (RF) classifiers. The confusion matrices obtained for the 3 elicitation modalities (ie, static, coherent, PC) are reported considering only the KNN and the RF classifiers.



Discussion

Principal Findings

The results confirm RQ1 ("A humanoid robot interacting with coherent emotional behavior is rated higher in terms of emotional valence, arousal, and dominance compared to the web application") because the COH robot is rated significantly higher for all SAM dimensions (except positive arousal) with respect to the PC condition (Figure 3). However, it is worth noting that when speaking of negative elicitation, receiving a higher rating of valence means that the stimulus with the COH condition was perceived less negatively than the ones elicited with the others. RQ2 ("There are significant differences in terms of emotional valence, arousal, and dominance between the static robot compared to the robot that shows movement") is confirmed for the 3 dimensions for negative and neutral emotions (Figure 3). It is worth noting that these results confirm that the robot's movements cause the negative emotion to be perceived as less negative (STA valence median value=3.32; COH valence median value=5.13). As for the positive emotion, there were no significant differences, which could suggest that the robot's behavior per se did not affect the perception of the positive emotion.

The presented results did not confirm the RQ3 ("The embodiment of humanoid robot will not affect the emotion perception compared to the web application") for all elicited emotion and SAM constructs. Indeed, there were no significant differences between the STA and the PC elicitation for valence and arousal measured during negative and neutral elicitation (Figure 3). Conversely, COH and STA differed significantly from PC in terms of positive elicitation. These results suggest that robot embodiment per se has a role in the perception of dominance associated with negative and neutral emotions with respect to a standard web interface. On the contrary, as for the

XSL•FO
RenderX

positive emotion, embodiment seems to play a key role because both COH and STA elicitations differ from the web application in terms of valence and dominance.

The ability to recognize user emotions is a fundamental step in the development of socially aware robots (RQ4). The emotions were recognized with an average accuracy >0.88 over the 3 elicitation conditions. In addition, the amount of gaze also depends on the interpersonal dynamics between the partner and their personalities and on the intent of using gaze to communicate their internal state. Therefore, it is important to measure it during interactions. As shown in Table 2, the accuracy of COH stimulation was higher than that of the other 2 methods. In addition, the results in the confusion matrices were aligned with the perceived emotion (Figure 5). According to the SAM results, the valence ratings for positive elicitation elicited with PC were significantly different from the other 2 with lower median values. This trend is reflected in the confusion matrices obtained using RF classifiers.

Comparison With Prior Work

Previous qualitative studies have pointed out how incoherent behavior can generate hilarious reactions in humans [16]. The presented results suggest that we can observe something similar, even if the stimulus is coherent. It appears that the robot's behavior somewhat distracts from perceiving negative emotions, even if the behavior is aligned with the shown emotion. In addition, as confirmation, positive emotion was perceived significantly more positively than the PC modality, suggesting that robot movements make the robot more positive. Consequently, these results suggest that it is important to tailor the reaction of the robot appropriately to elicit a specific emotion. Indeed, if we need to stimulate—for a certain reason—negatively the users, we need to reduce the robot's body expression because they can decrease the perception of negative emotions. Alternatively, if we need to provide positive feedback to users, the combined actions of both verbal and nonverbal communication can be used.

A previous study [36] compared robots and web applications that focused on investigating preferences and acceptance, and they did not find any significant deviation in the quantitative results. In contrast, in this study, we focus on human emotion perception, and this perception seems to be influenced or biased by the emotion itself and the robot's movement. This finding highlights the significance of not just robot embodiment but also its cospeech gestures in designing social agents, particularly when evaluating all dimensions of emotions. Methodologically, the presented findings carry significant implications for the design of experimental protocols. Evaluating HRI cannot rely solely on videos, as they overlook the importance of physical interaction. In the literature, some papers [33] provide a user impression without direct interaction with a robot; the collected results can be biased because the participant missed the contribution in the perception related to embodiment. Take, for instance, the scenario where you are testing a new game application or software on a tablet meant for eventual integration into a robot. Particularly when assessing emotions, it's crucial to approach the generalization of results with caution. In this sense, the result could be altered because the emotions elicited could not be directly applicable when interacting with an embodied agent.

The results obtained for the STA robots with the KNN and RF classifiers were slightly improved with respect to the results obtained in our previous work [42] (average accuracy was equal to 0.85 with KNN and 0.98 with RF), where we used them in combination with encoders. It is also worth noting that after the feature selection process, only the features related to gaze were retained in the analysis. Gaze is extremely important in managing interpersonal interaction and also during human-robot conversation; indeed, it can be correlated with user engagement during conversation or mutual tasks [50,51].

Limitations of the Study

The limitations of this study were mainly related to the cohort of recruited participants. First, both cognitive and cultural backgrounds are factors that can influence the perception of emotions [52]. Some neurological pathologies (eg, Parkinson disease) can affect facial expressions, whereas others can affect body gestures and language (eg, autism spectrum disorders and apathy); consequently, emotion recognition accuracy in such cases can change. The RQs do not focus on investigating their role in emotion perception; consequently, we recruited cohorts of people comparable for cultural background and cognitive status to limit the impact of these factors. The second limitation of this study refers to how the emotion is evaluated; in this study, we evaluated each SAM dimension separately. The third limitation of this study relies on the supervised machine learning techniques used. In this study, we rely on standard supervised methods because our main RQs are not focused on learning methods; therefore, we apply the most used techniques.

Future Directions

In this context, by applying the findings and implications of this paper in the health care context, we can conclude that it is important to tailor the reaction of the robot properly; indeed, if we need to stimulate—for a certain clinical reason—the users negatively, we need to reduce the robot's body expression because they can decrease the perception of negative emotions. Alternatively, if we need to give positive feedback to the users, for instance, during an exercise, we can use the combined action of both verbal and nonverbal communication. To overcome the limitations of this study, future research can be planned to extend the study to include a different group of participants with some cognitive and physical disorders and different cultural backgrounds to evaluate the effect of these factors on emotion perceptions. Future studies should also investigate whether there are differences in combining valence-arousal domains, as proposed in other studies [16,53]. Finally, the data could be analyzed using also deep learning and reinforcement learning techniques.

Conclusions

This study aimed to investigate the role of robot embodiment and its behavior in emotion perception and recognition using a standard elicitation model. In total, 4 RQs were investigated to understand how the robot's nonverbal behavior might shape the human perception of the showed emotion elicited through a standard data set and, at the same time, to observe the robot's

XSL•FO
RenderX

emotion recognition accuracy rate. This study presents an experimental setup in which 60 participants were asked to interact with 2 embodied agents (ie, a robot or tablet) that acted as emotion facilitators by showing them 60 standard pictures. The results underline the good recognition accuracy of the perception modules of the robot. Indeed, we can correctly classify the valence of the emotion (ie, positive, neutral, and negative) with an accuracy of up to 0.97 in the best case. According to the results, robot embodiment affects the perception of dominance significantly compared with web applications, which means that participants' emotions were less controlled when they were interacting with an embodied agent.

Acknowledgments

The authors would like to thank all the people involved in the study. This study was funded by "An adapted behavioral robot model with advanced cognitive/physical interaction capabilities for assessment and rehabilitation of neurodegenerative diseases (DESTINI)" founded by Unione Europea—NextGenerationEU (CUP: B55F21007810001).

Data Availability

The data sets generated and analyzed during this study are not publicly available because of the scope of the consent signed by the patient participating in the study but are in part (no video recordings) available from the corresponding author on reasonable request.

Authors' Contributions

The conceptualization was done by LF, GDO, and F Cavallo. Data curation was conducted by GDO, F Ciccone and AS. LF, FGCL, and AS were responsible for the data analysis. LF acquired the funding. The methodology was developed by GDO, LF, F Cavallo, and FG. The investigation was carried out by GDO and F Ciccone. AS and SR handled the software. DS, FG, and F Cavallo provided supervision. LF was responsible for the original draft of writing, while all authors contributed to the writing, review, and editing.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Self-Assessment Manikin questionnaire.
[PDF File (Adobe PDF File), 134 KB - humanfactors_v11i1e45494_app1.pdf]

Multimedia Appendix 2
Emotion recognition analysis.
[PDF File (Adobe PDF File), 194 KB - humanfactors_v11i1e45494_app2.pdf]

Multimedia Appendix 3
Participants' description.
[PDF File (Adobe PDF File), 120 KB - humanfactors_v11i1e45494_app3.pdf]

Multimedia Appendix 4
Median and IQR values computed for each elicited emotion.
[PDF File (Adobe PDF File), 143 KB - humanfactors_v11i1e45494_app4.pdf]

References

1. Cavallo F, Esposito R, Limosani R, Manzi A, Bevilacqua R, Felici E, et al. Robotic services acceptance in smart environments with older adults: user satisfaction and acceptability study. J Med Internet Res 2018 Oct 21;20(9):e264 [FREE Full text] [doi: 10.2196/jmir.9460] [Medline: 30249588]
2. Mannion A, Summerville S, Barrett E, Burke M, Santorelli A, Kruschke C, et al. Introducing the social robot MARIO to people living with dementia in long term residential care: reflections. Int J Soc Robot 2019 Jun 14;12(2):535-547. [doi: 10.1007/s12369-019-00568-1]
3. Pfadenhauer M, Dukat C. Robot caregiver or robot-supported caregiving? Int J Soc Robot 2015 Jan 30;7(3):393-406. [doi: 10.1007/s12369-015-0284-0]
4. Dawe J, Sutherland C, Barco A, Broadbent E. Can social robots help children in healthcare contexts? A scoping review. BMJ Paediatr Open 2019 Jan 31;3(1):e000371 [FREE Full text] [doi: 10.1136/bmjpo-2018-000371] [Medline: 30815587]

XSL•FO
RenderX

5. Rasouli S, Gupta G, Nilsen E, Dautenhahn K. Potential applications of social robots in robot-assisted interventions for social anxiety. Int J Soc Robot 2022 Jan 25;14(5):1-32 [FREE Full text] [doi: 10.1007/s12369-021-00851-0] [Medline: 35096198]
6. Scoglio AA, Reilly ED, Gorman JA, Drebing CE. Use of social robots in mental health and well-being research: systematic review. J Med Internet Res 2019 Jul 24;21(7):e13322 [FREE Full text] [doi: 10.2196/13322] [Medline: 31342908]
7. Hung L, Liu C, Woldum E, Au-Yeung A, Berndt A, Wallsworth C, et al. The benefits of and barriers to using a social robot PARO in care settings: a scoping review. BMC Geriatr 2019 Aug 23;19(1):232 [FREE Full text] [doi: 10.1186/s12877-019-1244-6] [Medline: 31443636]
8. Rossi S, Santini SJ, Di Genova D, Maggi G, Verrotti A, Farello G, et al. Using the social robot NAO for emotional support to children at a pediatric emergency department: randomized clinical trial. J Med Internet Res 2022 Jan 13;24(1):e29656 [FREE Full text] [doi: 10.2196/29656] [Medline: 34854814]
9. Henschel A, Laban G, Cross ES. What makes a robot social? A review of social robots from science fiction to a home or hospital near you. Curr Robot Rep 2021 Feb 04;2(1):9-19 [FREE Full text] [doi: 10.1007/s43154-020-00035-0] [Medline: 34977592]
10. Onyeulo EB, Gandhi V. What makes a social robot good at interacting with humans? Information 2020 Jan 13;11(1):43. [doi: 10.3390/info11010043]
11. de Graaf MM, Ben Allouch S, van Dijk JA. What makes robots social?: a user's perspective on characteristics for social human-robot interaction. In: Proceedings of the Social Robotics. 2015 Presented at: ICSR 2015; October 26-30, 2015; Paris, France. [doi: 10.1007/978-3-319-25554-5_19]
12. Horstmann AC, Krämer NC. Great expectations? Relation of previous experiences with social robots in real life or in the media and expectancies based on qualitative and quantitative assessment. Front Psychol 2019 Apr 30;10:939 [FREE Full text] [doi: 10.3389/fpsyg.2019.00939] [Medline: 31114525]
13. Rosenthal-von der Pütten AM, Krämer NC, Herrmann J. The effects of humanlike and robot-specific affective nonverbal behavior on perception, emotion, and behavior. Int J Soc Robot 2018 Jan 23;10(5):569-582. [doi: 10.1007/s12369-018-0466-7]
14. Mavridis N. A review of verbal and non-verbal human–robot interactive communication. Robot Auton Syst 2015 Jan;63:22-35. [doi: 10.1016/j.robot.2014.09.031]
15. Rossi S, Ruocco M. Better alone than in bad company. Interact Stud 2019 Nov 18;20(3):487-508. [doi: 10.1075/is.18066.ros]
16. Rossi S, Cimmino T, Matarese M, Raiano M. Coherent and incoherent robot emotional behavior for humorous and engaging recommendations. In: Proceedings of the 28th IEEE International Conference on Robot and Human Interactive Communication (RO-MAN). 2019 Presented at: 28th IEEE International Conference on Robot and Human Interactive Communication (RO-MAN); October 14-18, 2019; New Delhi, India. [doi: 10.1109/ro-man46459.2019.8956260]
17. Canamero L, Fredslund J. I show you how I like you - can you read it in my face? [robotics]. IEEE Trans Syst Man Cybern Part A Syst Humans 2001 Sep;31(5):454-459. [doi: 10.1109/3468.952719]
18. Tsiourti C, Weiss A, Wac K, Vincze M. Multimodal integration of emotional signals from voice, body, and context: effects of (in)congruence on emotion recognition and attitudes towards robots. Int J Soc Robot 2019 Feb 4;11(4):555-573. [doi: 10.1007/s12369-019-00524-z]
19. Sorrentino A, Khalid O, Coviello L, Cavallo F, Fiorini L. Modeling human-like robot personalities as a key to foster socially aware navigation. In: Proceedings of the 30th IEEE International Conference on Robot & Human Interactive Communication (RO-MAN). 2021 Presented at: 30th IEEE International Conference on Robot & Human Interactive Communication (RO-MAN); August 08-12, 2021; Vancouver, BC. [doi: 10.1109/ro-man50785.2021.9515556]
20. Guo F, Li M, Qu Q, Duffy VG. The effect of a humanoid robot's emotional behaviors on users' emotional responses: evidence from pupillometry and electroencephalography measures. Int J Hum Comput Interact 2019 Mar 29;35(20):1947-1959. [doi: 10.1080/10447318.2019.1587938]
21. Umbrico A, Cesta A, Cortellessa G, Orlandini A. A holistic approach to behavior adaptation for socially assistive robots. Int J Soc Robot 2020 Jan 28;12(3):617-637. [doi: 10.1007/s12369-019-00617-9]
22. Spezialetti M, Placidi G, Rossi S. Emotion recognition for human-robot interaction: recent advances and future perspectives. Front Robot AI 2020 Dec 21;7:532279 [FREE Full text] [doi: 10.3389/frobt.2020.532279] [Medline: 33501307]
23. Yan F, Iliyasu AM, Hirota K. Emotion space modelling for social robots. Eng Appl Artif Intell 2021 Apr;100:104178. [doi: 10.1016/j.engappai.2021.104178]
24. Cavallo F, Semeraro F, Fiorini L, Magyar G, Sinčák P, Dario P. Emotion modelling for social robotics applications: a review. J Bionic Eng 2018 Mar 23;15(2):185-203. [doi: 10.1007/s42235-018-0015-y]
25. Fiorini L, Loizzo FG, Sorrentino A, Kim J, Rovini E, Di Nuovo A, et al. Daily gesture recognition during human-robot interaction combining vision and wearable systems. IEEE Sensors J 2021 Oct 15;21(20):23568-23577. [doi: 10.1109/jsen.2021.3108011]
26. Manzi A, Fiorini L, Limosani R, Dario P, Cavallo F. Two - person activity recognition using skeleton data. IET Comput Vis 2017 Oct 20;12(1):27-35. [doi: 10.1049/iet-cvi.2017.0118]
27. Albanie S, Nagrani A, Vedaldi A, Zisserman A. Emotion recognition in speech using cross-modal transfer in the wild. In: Proceedings of the 26th ACM international conference on Multimedia. 2018 Presented at: MM '18; October 22-26, 2018; Seoul, Republic of Korea. [doi: 10.1145/3240508.3240578]

XSL•FO
RenderX

28. Juszkiewicz L. Improving speech emotion recognition system for a social robot with speaker recognition. In: Proceedings of the 19th International Conference on Methods and Models in Automation and Robotics (MMAR). 2014 Presented at: 19th International Conference on Methods and Models in Automation and Robotics (MMAR); September 02-05, 2014; Miedzyzdroje, Poland. [doi: 10.1109/mmar.2014.6957480]
29. Anzalone SM, Boucenna S, Ivaldi S, Chetouani M. Evaluating the engagement with social robots. Int J Soc Robot 2015 Apr 17;7(4):465-478. [doi: 10.1007/s12369-015-0298-7]
30. Nocentini O, Fiorini L, Acerbi G, Sorrentino A, Mancioppi G, Cavallo F. A survey of behavioral models for social robots. Robotics 2019 Jul 09;8(3):54. [doi: 10.3390/robotics8030054]
31. Rossi S, Ferland F, Tapus A. User profiling and behavioral adaptation for HRI: a survey. Pattern Recognit Lett 2017 Nov;99:3-12. [doi: 10.1016/j.patrec.2017.06.002]
32. Bradley MM, Lang PJ. Measuring emotion: the self-assessment manikin and the semantic differential. J Behav Ther Exp Psychiatry 1994 Mar;25(1):49-59. [doi: 10.1016/0005-7916(94)90063-9] [Medline: 7962581]
33. Barchard KA, Lapping-Carr L, Westfall RS, Fink-Armold A, Banisetty SB, Feil-Seifer D. Measuring the perceived social intelligence of robots. ACM Trans Hum Robot Interact 2020 Sep 11;9(4):1-29. [doi: 10.1145/3415139]
34. Bagheri E, Esteban PG, Cao HL, De Beir A, Lefeber D, Vanderborght B. An autonomous cognitive empathy model responsive to users' facial emotion expressions. ACM Trans Interact Intell Syst 2020 Nov 09;10(3):1-23. [doi: 10.1145/3341198]
35. Rossi S, Staffa M, Tamburro A. Socially assistive robot for providing recommendations: comparing a humanoid robot with a mobile application. Int J Soc Robot 2018 Jan 31;10(2):265-278. [doi: 10.1007/s12369-018-0469-4]
36. Staffa M, Rossi S. Recommender interfaces: the more human-like, the more humans like. In: Proceedings of the Social Robotics. 2016 Presented at: ICSR 2016; November 1-3, 2016; Kansas City, MO. [doi: 10.1007/978-3-319-47437-3_20]
37. Becker C, Prendinger H, Ishizuka M, Wachsmuth I. Evaluating affective feedback of the 3D agent max in a competitive cards game. In: Proceedings of the Affective Computing and Intelligent Interaction. 2005 Presented at: ACII 2005; October 22-24, 2005; Beijing, China URL: https://link.springer.com/chapter/10.1007/11573548_60 [doi: 10.1007/11573548_60]
38. Lang PJ, Bradley MM, Cuthbert BN. International affective picture system (IAPS): affective ratings of pictures and instruction manual. Technical report A-8. University of Florida. 2008. URL: https://gitlab.pavlovia.org/rsaitov/experimental-psycholoy-ltu-final/raw/563e2b213be7eb54829116f05b83f0124362d71b/IAPS.TechManual.1-20.2008(1).pdf [accessed 2020-02-01]
39. Russell JA, Mehrabian A. Evidence for a three-factor theory of emotions. J Res Pers 1977 Sep;11(3):273-279 [FREE Full text] [doi: 10.1016/0092-6566(77)90037-X]
40. Salem M, Rohlfing K, Kopp S, Joublin F. A friendly gesture: investigating the effect of multimodal robot behavior in human-robot interaction. In: Proceedings of the 2011 RO-MAN. 2011 Presented at: 2011 RO-MAN; July 31-August 03, 2011; Atlanta, GA. [doi: 10.1109/roman.2011.6005285]
41. Fiorini LF, Loizzo FG, D'Onofrio G, Sorrentino A, Ciccone F, Russo S, et al. Can I feel you? Recognizing human's emotions during human-robot interaction. In: Proceedings of the Social Robotics. 2022 Presented at: ICSR 2022; December 13-16, 2022; Florence, Italy. [doi: 10.1007/978-3-031-24667-8_45]
42. D'Onofrio G, Fiorini L, Sorrentino A, Russo S, Ciccone F, Giuliani F, et al. Emotion recognizing by a robotic solution initiative (EMOTIVE project). Sensors (Basel) 2022 May 08;22(8):2861 [FREE Full text] [doi: 10.3390/s22082861] [Medline: 35458845]
43. Social robot. Behaviour Labs. URL: https://www.blabs.eu/robomate [accessed 2023-03-30]
44. Lang PJ, Bradley MM, Cuthbert BN. International affective picture system (IAPS): technical manual and affective ratings. Federal University of São Paulo. 1997. URL: https://www2.unifesp.br/dpsicobio/adap/instructions.pdf [accessed 2023-12-31]
45. Hall ET. The Hidden Dimension. New York, NY: Knopf Doubleday Publishing Group; 1966.
46. Gatti E, Calzolari E, Maggioni E, Obrist M. Emotional ratings and skin conductance response to visual, auditory and haptic stimuli. Sci Data 2018 Jul 26;5(1):180120 [FREE Full text] [doi: 10.1038/sdata.2018.120] [Medline: 29944144]
47. Avidemux. Uptodown. URL: https://avidemux.it.uptodown.com/windows [accessed 2023-03-30]
48. Baltrusaitis T, Zadeh A, Lim YC, Morency LP. OpenFace 2.0: facial behavior analysis toolkit. In: Proceedings of the 13th IEEE International Conference on Automatic Face & Gesture Recognition (FG 2018). 2018 Presented at: 13th IEEE International Conference on Automatic Face & Gesture Recognition (FG 2018); May 15-19, 2018; Xi'an, China. [doi: 10.1109/fg.2018.00019]
49. Lara OD, Labrador MA. A survey on human activity recognition using wearable sensors. IEEE Commun Surv Tutor 2012 Nov 29;15(3):1192-1209. [doi: 10.1109/SURV.2012.110112.00192]
50. Admoni H, Scassellati B. Social eye gaze in human-robot interaction: a review. J Hum Robot Interact 2017 May;6(1):25-63. [doi: 10.5898/JHRI.6.1.Admoni]
51. Kompatsiari K, Bossi F, Wykowska A. Eye contact during joint attention with a humanoid robot modulates oscillatory brain activity. Soc Cognit Affect Neurosci 2021;16(4):383-392. [doi: 10.31234/osf.io/eam3d]
52. Fiorini L, Coviello L, Sorrentino A, Sancarlo D, Ciccone F, D'Onofrio G, et al. User profiling to enhance clinical assessment and human-robot interaction: a feasibility study. Int J Soc Robot 2023;15(3):501-516 [FREE Full text] [doi: 10.1007/s12369-022-00901-1] [Medline: 35846164]

XSL•FO
RenderX

53. Fiorini L, Mancioppi G, Semeraro F, Fujita H, Cavallo F. Unsupervised emotional state classification through physiological parameters for social robotics applications. Knowl Based Syst 2020 Feb;190:105217. [doi: 10.1016/j.knosys.2019.105217]

Abbreviations

COH: coherent
HRI: human-robot interaction
IAPS: International Affective Picture System
KNN: K-nearest neighbor
RF: random forest
RQ: research question
SAM: Self-Assessment Manikin questionnaire
STA: static
SVM: support vector machine

Edited by A Choudhury; submitted 04.01.23; peer-reviewed by O Bendel, M Amith; comments to author 21.03.23; revised version received 24.04.23; accepted 29.11.23; published 26.01.24.

Please cite as:
Fiorini L, D'Onofrio G, Sorrentino A, Cornacchia Loizzo FG, Russo S, Ciccone F, Giuliani F, Sancarlo D, Cavallo F
The Role of Coherent Robot Behavior and Embodiment in Emotion Perception and Recognition During Human-Robot Interaction: Experimental Study
JMIR Hum Factors 2024;11:e45494
URL: https://humanfactors.jmir.org/2024/1/e45494
doi:10.2196/45494
PMID:38277201

©Laura Fiorini, Grazia D'Onofrio, Alessandra Sorrentino, Federica Gabriella Cornacchia Loizzo, Sergio Russo, Filomena Ciccone, Francesco Giuliani, Daniele Sancarlo, Filippo Cavallo. Originally published in JMIR Human Factors (https://humanfactors.jmir.org), 26.01.2024. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on https://humanfactors.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

The Effect of a Video-Assisted Health Education Program Followed by Peer Education on the Health Literacy of COVID-19 and Other Infectious Diseases Among School Children: Quasi-Randomized Controlled Trial

Xiaojuan Zhang[1,2], BMed; Yingkun Justin Wen[1,3], BMed; Ning Han[4*], BMed; Yawen Jiang[1*], PhD

[1]School of Public Health (Shenzhen), Sun Yat-sen University, Shenzhen, Guangdong, China

[2]School of Public Health, Sun Yat-sen University, Guangzhou, China

[3]Department of Learning, Informatics, Management and Ethics, Karolinska Institutet, Stockholm, Sweden

[4]Institute of Public Health Supervision of Longgang District, Shenzhen, China

[*]these authors contributed equally

Corresponding Author:
Yawen Jiang, PhD
School of Public Health (Shenzhen)
Sun Yat-sen University
Room 533, West Wing of Medical Complex #1, Sun Yat-sen University
66 Gongchang Road, Guangming District Shenzhen
Shenzhen, Guangdong
China
Phone: 86 13632974660
Email: jiangyw26@mail.sysu.edu.cn

Abstract

Background: To improve the engagement and effectiveness of traditional health programs, it is necessary to explore alternative models of health education including video-assisted lectures and peer education.

Objective: This study aimed to evaluate the effects of a combination of video-assisted lectures and peer education on health literacy related to infectious diseases among students.

Methods: Third-grade classes from 11 pilot schools in Longgang District of Shenzhen, China, were randomized to the intervention and control groups. In the intervention group, a video-assisted interactive health education program was conducted twice over a time span of 5 months. Each of the 2 sessions included a 40-minute lecture on COVID-19 and other common infectious diseases in schools and a 5-minute science video. In addition, 5 “little health supervisors” at the end of the first session were elected in each class, who were responsible for helping class members to learn health knowledge and develop good hygiene habits. Students answered the same quiz before the first and after the second session. Models based on item response theory (IRT) were constructed to score the students’ knowledge of infectious diseases based on the quiz.

Results: In total, 52 classes and 2526 students (intervention group: n=1311; control group: n=1215) were enrolled. Responses of the baseline survey were available for 2177 (86.2%; intervention group: n=1306; control group: n=871) students and those of the postintervention survey were available for 1862 (73.7%; intervention group: n=1187; control group: n=675). There were significant cross-group differences in the rates of correctly answering questions about influenza symptoms, transmission, and preventive measures; chicken pox symptoms; norovirus diarrhea symptoms; mumps symptoms; and COVID-19 symptoms. Average IRT scores of questions related to infectious diseases in the intervention and control groups were, respectively, –0.0375 (SD 0.7784) and 0.0477 (SD 0.7481) before the intervention (P=.01), suggesting better baseline knowledge in the control group. After the intervention, the average scores of the intervention and control groups were 0.0543 (SD 0.7569) and –0.1115 (SD 0.7307), respectively (P<.001), suggesting not only significantly better scores but also greater improvement in the intervention group.

Conclusions: After the health education project, the correct answer rate of infectious disease questions in the intervention group was higher than that of the control group, which indicates significant effects of the combination of video-assisted lectures and peer education for the promotion of health literacy. In addition, the intervention effect of the first session persisted for at least 4

XSL•FO
RenderX

months up to the second session. As such, the proposed program was effective in improving the health literacy of school children in relation to infectious diseases and should be considered for massive health promotion campaigns during pandemics.

Trial Registration: ISRCTN ISRCTN49297995; https://www.isrctn.com/ISRCTN49297995

(JMIR Hum Factors 2024;11:e43943) doi:10.2196/43943

KEYWORDS

infectious diseases; primary school students; quasi-randomized controlled trial; video-assisted health education; peer education; item response theory; IRT

Introduction

Primary school students are vulnerable to emerging and common infectious diseases such as COVID-19, influenza, mumps, and intestinal infectious diseases [1]. A survey on the reasons for sick leaves in primary and secondary schools in Shenzhen showed that the top 5 causes were common cold, gastrointestinal diseases, unexplained or other illness, influenza, and chicken pox [2]. In addition, the importance of acquiring essential knowledge regarding the prevention and control of COVID-19 cannot be overstated during the pandemic. Accordingly, it is critical to embed health promotion into the school education of primary school students. To that end, the outline of "Healthy China 2030" emphasizes the importance of fortifying health education among school children. In particular, primary schools were integral to the life cycle of the health education curriculum to the extent that early-life exposure to information on diseases and health behaviors is associated with improved future health outcomes [3].

Despite its importance, health education was highly restricted in its delivery forms. Conventionally, the most prevalent approach of health education of infectious diseases for school students was, arguably, classroom lectures aided with paper-based materials, in which the teaching contents are usually compiled by school teachers and researchers [4]. Traditional health education is also reported to have a limited duration of effects. Hampered by the collective challenges faced in traditional health education, most schools lack systematic health education programs [5]. To increase students' interest in healthy behaviors and to extend the duration of education effects, researchers have been exploring alternative media for health education. Among the various new models, two of the prevailing approaches are video-assisted health education and interactional peer education [5,6].

In professional medical education, video-assisted lectures are useful tools for students to acquire basic clinical skills. When delivered in bundle with in-person lectures, video-based materials are often preferred by students [7]. In addition, video-assisted health education has been shown to be more effective than oral education in facilitating postoperative recovery of patients [8].

The effects of health education are not necessarily limited to the immediate recipients of the program themselves. Students may also help to shape the opinions and behaviors of their classmates by becoming peer educators of health and hygiene. Peer education is defined as "sharing experiences and learning among people with something in common," such as a similar age, living environment, and culture [9]. There is substantial evidence that peer education is highly effective in specific areas of medical and health education, including professional medical training, chronic disease prevention, and sexual health behaviors [6,10,11]. Incorporating peer effects into the design of health education programs could, therefore, strengthen the programs' impacts on behavioral change.

However, evidence on the effects of video-assisted lectures and peer education on health literacy among school children is still lacking. Given its substantial potential for public health practice, we designed a health education package that combined video-assisted classroom teaching and peer education and tested the effectiveness of this program. This program, which we anecdotally refer to as the "Little Health Supervisors" project, was anticipated to improve the health literacy of students over an array of infectious diseases.

Methods

Trial Design

The "Little Health Supervisors" project is jointly enacted by the Longgang District Bureau of Health and the Longgang District Bureau of Education as an administrative task. Third-grade classes from 11 pilot schools in Longgang district of Shenzhen, China, were randomized to the intervention and control groups. Our aim was to allocate equal numbers of third-grade classrooms to the intervention and control groups within each school. However, schools with an odd total number of classes inevitably resulted in uneven groups; hence, one group might outnumber another eventually. This project enclosed 2 health education sessions 4 months apart in Dec 2021 and Apr 2022 in Longgang District, Shenzhen City in the Guangdong Province of China, which is a district with approximately 4 million residents and 0.4 million school students.

Ethical Considerations

The "Little Health Supervisors" project was launched by the district government as a public service project. The study protocol was approved by the Biomedical Research Ethics Review Committee, School of Public Health (Shenzhen), Sun Yat-sen University [2021(056)] and was registered with Longgang District Bureau of Health (Figure S1 in Multimedia Appendix 1). Informed consent was obtained from all students and their parents who met the inclusion criteria and were willing to participate. Confidentiality of information was maintained.

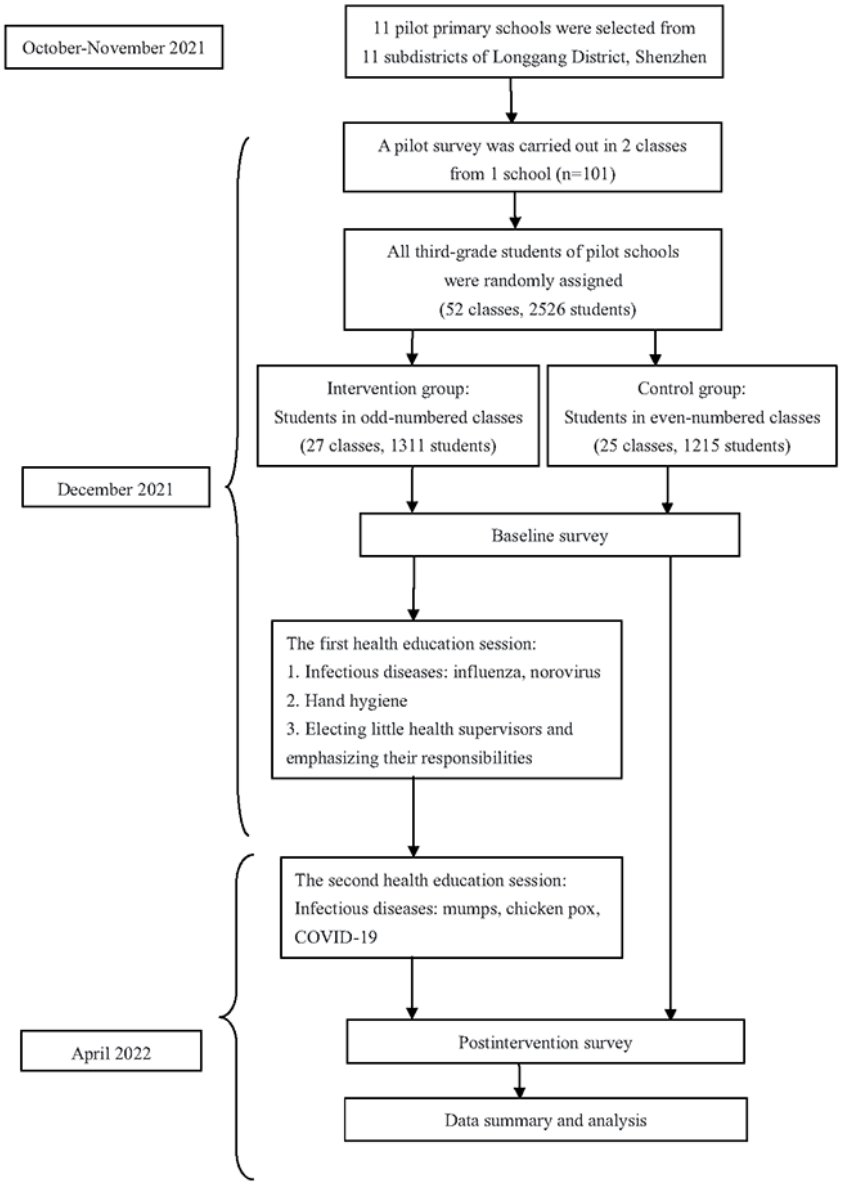
Recruitment

In the first step of sample enrollment, considering the feasibility of the project's implementation, the Longgang District Bureau

XSL•FO
RenderX

of Health and the Longgang District Bureau of Education recommended 1 primary school based on the willingness to participate for each of the 11 subdistricts of this district. Second, all third-grade students in the 11 schools were eligible for participation if they met the following requirements: (1) they were not taking a leave of absence from school at the time of enrollment; (2) they agreed (or their guardians agreed) to spend time on attending lectures; (3) they had access to a computer, tablet, or smartphone with an internet connection; (4) they had sufficient knowledge to use mobile devices or computers (assistance allowed); and (5) they were able to read and interpret Chinese characters. Next, as decided by the researchers, the eligible students were assigned to the intervention and control groups using the class number as the randomizer. Specifically, odd-numbered classes were assigned to the intervention group; even-numbered classes, the control group. The flowchart of participant enrollment is illustrated in Figure 1.

Figure 1. Flowchart of the "Little Health Supervisors" project (a cluster randomized controlled trial) from December 2021 to April 2022.



Data Collection

To standardize students' knowledge of infectious diseases both before and after the education program, a questionnaire containing a quiz on COVID-19 and selected infectious diseases with relatively high local incidences was curated, which included influenza, chicken pox, norovirus diarrhea, and mumps. The questionnaire also collected demographic characteristics (school, class, student number, sex, and date of birth) and COVID-19 vaccination status. In addition, we delivered a separate questionnaire to a parent of each student who collected parental assent to vaccinate their children against COVID-19. Moreover, family socioeconomic information was also collected in the parent questionnaire, which included monthly household income and the parents' education level [12].

Questionnaires were distributed via a web-based survey platform (Wenjuanxing, Changsha Ranxing Information Technology Co, Ltd). In the baseline survey, students completed the questionnaires in a computer laboratory with the instructions of either the computer teachers or the class advisors. To collect

XSL•FO
RenderX

parents' responses, the teachers arranged a meeting with each family using previously connected social media to select a representative for questionnaire responses. Due to COVID-19 outbreaks during the planned time period of the second session, the postintervention survey was distributed on the web. Simultaneously, the researchers also collected the questionnaire from the control group.

Interventions

The intervention was developed by both researchers and the local health department. Details of the development process and the content of the intervention are provided in Table S1 and Figures S2-S3 in Multimedia Appendix 1 [13,14]. Students randomized to the intervention group had access to 2 free sessions of health education during the study. Each session included a 40-minute lecture on the transmission and prevention of different infectious diseases, followed by a 5-minute science video. To incentivize learning, students were informed that there would be interactional question-and-answer sections during the lecture, for which the participating students were eligible for prizes.

In December 2021, the baseline survey and the first health education session were conducted, with the former preceding the latter. The in-person lecture and the videos of the first session pertained to influenza, norovirus diarrhea, and hand hygiene. At the end of this session, 5 little health supervisors were elected by the teachers from each class. They were naturally assumed as opinion leaders, showcasing their ability to effectively convey knowledge and could supervise the learning of health knowledge and the development of good hygiene habits of their classmates. The teachers also handed out brochures, armbands, and stickers to the 5 little health supervisors. In addition, the teachers encouraged all students to take health knowledge home and improve the family's health literacy by way of "small hands holding big hands," which aimed to exploit the power of two-step flow theory of communication for information transmission. Originating from political science, the two-step flow theory asserts that information can be conveyed through the chain of media-opinion leaders-audience. Students may also help to shape the opinions and behaviors of their family members by becoming an opinion leader of health and hygiene [15,16].

In April 2022, the second health education session and the postintervention survey were carried out. However, the order of education and survey was reversed in relation to the first session. The lecture and the videos of the second session pertained to chicken pox, mumps, and COVID-19 symptoms. Affected by a local COVID-19 outbreak, students had to take the web-based classes at home, so the health education sessions had to be conducted in the form of recorded course videos. In the intervention group, students were required to watch the video, and the teachers also encouraged all students to distribute health knowledge to the people around them.

As for the control group, the students only received routine health education at school, which included health tips on influenza from school doctors and 1 or 2 public welfare courses conducted by the local health department or hospitals every semester. These routine health education sessions were balanced between the 2 groups.

Outcomes

The primary outcomes of this trial were the score in the original scale (hereafter referred to as "crude score") and item response theory (IRT) score of questions related to infectious diseases, the correct answer rates of questions related to infectious diseases, and the pre-post changes in the correct answer rates after the intervention. The secondary outcomes were the COVID-19 vaccination rates. For those who did not receive COVID-19 vaccines at baseline or at the end of the program, we also exploratively asked about their willingness to get vaccinated and the reasons for not being vaccinated.

Statistical Analysis

To gain an overview of students' characteristics, their families' demographic data were collected. Monthly household income (in ¥) was categorized into 4 levels (<¥5000 [US $702.97], ¥5000 [US $702.97]~¥10,000 [US $1405.94], ¥10,000 [US $1405.94]~¥20,000 [US $2811.88], and ≥¥20,000 [US $2811.88]). Parent's education was grouped into 3 levels (junior high or below, secondary school [including technical secondary school], and college and above]. For the questions related to infectious diseases, multiple answers were regarded as correct only if all the correct answers were selected. Correctly answered questions contributed 1 point, and incorrectly answered questions contributed 0 points. The crude score of questions related to infectious diseases ranged from 0 to 7, with a higher score indicating higher knowledge of infectious diseases. For the item of willingness to be vaccinated against COVID-19, we assigned 1, 2, 3, 4, and 5 points respectively to the 5 options of very reluctant, reluctant, neutral, willing, and very willing. To comprehensively evaluate the students' knowledge of infectious diseases, IRT was used to fit the model of 7 items of the questionnaire. Frequently used in studies on education examinations, IRT is a set of psychometric models used to measure unobservable characteristics of the respondents and the development of scoring scales [17-19]. IRT can be used to explain the relationship between a latent trait (eg, the health literacy of school children related to infectious diseases) and observable characteristics and items (eg, questionnaire answers). IRT has at least 3 model specifications. The one parameter logistic model takes item difficulty into account when evaluating individual ability, whereas the two parameter logistic model additionally considers differential discrimination of items [19,20]. In addition to these 2 models, the three-parameter model (TPM) allows the possibility of guessing [19,20]. In this study, a TPM was selected to calculate the IRT score (Table S2 in Multimedia Appendix 1). To score the students' latent health literacy, we fitted TPM using the R package "ltm: Birnbaum's three parameter model" to the 7 questions related to the knowledge of infectious diseases [20]. A higher score meant higher health literacy. We plotted the estimated IRT score of questions related to infectious diseases to visualize the students' performance (Figures S4-S7 in Multimedia Appendix 1). Although not directly related to our main analyses, we also plotted the item characteristic curves, item information curves, and the test information curve to provide some information

XSL•FO
RenderX

regarding the difficulty of the test (Figures S8-S10 in Multimedia Appendix 1).

Finally, to summarize categorial sociodemographic characteristics, the correct answer rates of answering the questions, the pre-post changes in the correct answer rates after the intervention, the COVID-19 vaccination rate, the reasons for nonvaccination, and the percentages of the corresponding variables were calculated. We used mean and SD to describe the crude score, the IRT score of questions related to infectious diseases, and the willingness to vaccinate against COVID-19. We used *t* tests to compare the crude score, the IRT score, and the willingness to be vaccinated against COVID-19 across groups. Regarding the willingness to be vaccinated between 2 groups, we also conducted a stratified analysis based on the parents' sex. Chi-square tests were carried out on the basis of the correct answer rate, the COVID-19 vaccination rate, and the reason for nonvaccination to investigate differences between the 2 groups. The pre-post changes in the correct answer rates after the intervention were compared between study groups, using the *z* test. Furthermore, since we used class as our intervention unit, we also conducted an additional analysis using class as the primary unit of analysis. This was undertaken to ensure that our class-based examination would yield coherent findings as well (Tables S3-S5 in Multimedia Appendix 1). A *P* value less than .05 was considered significant. All data were analyzed using SPSS (version 26; IBM Corp) and R (version 4.2.0; The R Foundation).

Power

We calculated the power of this study on the basis of the sample size of the intervention and on the primary outcome. To calculate power, we used the sample size of 1862 (intervention group: n=1187; control group: n=675), an acceptable probability for type I error of .05, a pooled SD of 0.767, and a minimal difference in the infectious disease knowledge scores between the 2 groups of 0.166 (ie, μ_1–μ_2). The power of this study was 99.43%.

Data Exclusion

First, when an intervention group student decided to quit or was lost to follow-up, the student was excluded from the primary analysis. Second, the researchers checked information such as IP address, birth date, sex, and school and class codes to identify duplicates.

Results

Study Population

In the baseline survey, 2177 (intervention group: n=1306; control group: n=871) student questionnaires and 2496 (intervention group: n=1430; control group: n=1066) parent questionnaires were collected, amounting to response rates of 86.2% and 98.8%, respectively. In the postintervention survey, 1862 (intervention group: n=1187; control group: n=675) student questionnaires and 1799 (intervention group: n=1076; control group: n=723) parent questionnaires were retrieved, yielding response rates of 73.7% and 71.2%, respectively (Tables S6-S9 in Multimedia Appendix 1). In the intervention group, 2493 (intervention group: n=1306; control group: n=1187) student questionnaires were collected, with a response rate of 95.1%. In the control group, 1546 (baseline survey: n=871; postintervention survey: n=675) student questionnaires were collected, with a response rate of 63.6%.

There were no significant differences in baseline characteristics between the intervention and control groups (Table 1). In the intervention group, there were 691 male and 615 female students; the corresponding numbers in the control group were 459 and 412, respectively. The proportion of households earning less than ¥5000 (US $702.97) was relatively small in both groups (9.8% and 8.9%). Finally, the proportions of students whose parents had college education and above was 72.7% in both groups.

Table 1. Sociodemographic characteristics of third-grade students from 11 pilot schools in Longgang District of Shenzhen, China.

Characteristics	Intervention group, n/n (%)	Control group, n/n (%)
Sex		
Male	691/1306 (52.9)	459/871 (52.7)
Female	615/1306 (47.1)	412/871 (47.3)
Monthly household income (¥[a])		
<5000	140/1430 (9.8)	95/1066 (8.9)
5000~10,000	359/1430 (25.1)	267/1066 (25.0)
10,000~20,000	403/1430 (28.2)	311/1066 (29.2)
≥20,000	528/1430 (36.9)	393/1066 (36.9)
Parent's educational level		
Junior high or below	116/1430 (8.1)	82/1066 (7.7)
High school or technical secondary school	275/1430 (19.2)	209/1066 (19.6)
College and above	1039/1430 (72.7)	775/1066 (72.7)

[a]¥1=US $0.1445.

XSL•FO
RenderX

Correct Answer Rates of Questions Related to Infectious Diseases

At baseline, the correct answer rates for questions related to influenza symptoms, influenza preventive measures, and norovirus diarrhea symptoms were different between the intervention and control groups. Specifically, the correct answer rate was higher in the control group (Table 2). In terms of the correct answer rates for questions regarding influenza transmission, chicken pox symptoms, mumps transmission, and COVID-19 symptoms, there were no significant differences between the 2 groups (Table 2). After the intervention, the differences between the 2 groups in the correct answer rates for questions regarding influenza symptoms, influenza preventive measures, and norovirus diarrhea symptoms were no longer observed (Table 2). By contrast, the differences in the correct answer rates for questions regarding chicken pox symptoms, mumps transmission, and COVID-19 symptoms between the 2 groups at the end point were significant, such that intervention group outperformed the control group (Table 2).

Table 2. The correct answer rates for questions related to infectious diseases in the intervention and control groups.

Questions	Total, %	Intervention group, %	Control group, %	*P* value
Baseline				
Influenza symptoms	67.57	65.39	70.84	.008
Influenza transmission	81.86	81.47	82.43	.57
Influenza preventive measures	84.66	83.08	87.03	.01
Norovirus diarrhea symptoms	56.41	54.21	59.70	.01
Chicken pox symptoms	28.34	28.79	27.67	.57
Mumps transmission	6.89	7.27	6.31	.39
COVID-19 symptoms	29.54	28.33	31.34	.13
End point				
Influenza symptoms	86.09	86.77	84.89	.26
Influenza transmission	78.30	78.69	77.63	.60
Influenza preventive measures	92.91	93.09	92.59	.69
Norovirus diarrhea symptoms	72.93	73.80	71.41	.26
Chicken pox symptoms	43.18	47.01	36.44	<.001
Mumps transmission	10.15	13.23	4.74	<.001
COVID-19 symptoms	49.14	52.40	43.41	<.001

Regarding the pre-post changes in the correct answer rates after the intervention, the differences between the 2 groups were significant for all items (Table 3). Specifically, the correct answer rates for questions regarding influenza symptoms, influenza preventive measures, norovirus diarrhea symptoms, chicken pox symptoms, and COVID-19 symptoms increased in both groups (for all, $P<.001$). However, the correct answer rates of the intervention group increased more than those of the control group. In the intervention group, the correct answer rate for questions regarding mumps transmission increased in the intervention group but decreased slightly in the control group. Compared with that before the intervention, the correct answer rate for questions regarding influenza transmission decreased slightly after the intervention (Table 3).

Table 3. Pre-post changes in the correct answer rates after the intervention in the intervention and control groups.

Questions	Total, %	Intervention group, %	Control group, %	*P* value
Influenza symptoms	18.52	21.38	14.05	<.001
Influenza transmission	–3.56	–2.78	–4.80	.02
Influenza preventive measures	8.25	10.01	5.56	<.001
Norovirus diarrhea symptoms	16.52	19.59	11.71	<.001
Chicken pox symptoms	14.84	18.22	8.77	<.001
Mumps transmission	3.26	5.96	–1.57	<.001
COVID-19 symptoms	19.60	24.07	12.07	<.001

XSL•FO
RenderX

Crude and IRT Scores for Questions Related to Infectious Diseases

Before the intervention, there was a significant difference in the mean scores for questions regarding infectious disease knowledge between the 2 groups. The mean IRT score of the intervention group (–0.0375, SD 0.7784) was significantly lower (*P*=.01) than that of the control group (0.0477, SD 0.7481). After the intervention, the mean IRT score of the intervention group (0.0543, SD 0.7569) surpassed that of the control group (–0.1115, SD 0.7307). Notably, the postintervention mean score of the intervention group increased from that at baseline, whereas the control group displayed an opposite trend (Table 4). The situation is similar for the crude score (Table 4).

Table 4. The crude and item response theory (IRT) score of questions related to infectious diseases in the intervention and control groups.

	Crude score, mean (SD)			IRT-based score, mean (SD)		
	Intervention group	Control group	*P* value	Intervention group	Control group	*P* value
Baseline	3.49 (1.628)	3.65 (1.552)	.02	–0.0375 (0.7784)	0.0477 (0.7481)	.01
End point	4.45 (1.469)	4.11 (1.420)	<.001	0.0543 (0.7569)	–0.1115 (0.7307)	<.001

COVID-19 Vaccination Rates

The COVID-19 vaccination rates of the intervention and the control groups at baseline were 94.8% and 93.2%, respectively; by the end of the program, they increased slightly to 97.6% and 96.6%, respectively. The differences, however, were not significant (Table 5).

Table 5. The COVID-19 vaccination rates of third-grade students before and after the intervention.

	Intervention group, n/n (%)	Control group, n/n (%)	*P* value
Baseline	1238/1306 (94.8)	812/871 (93.2)	.13
End point	1158/1187 (97.6)	652/675 (96.6)	.23

Willingness to Get Vaccinated and the Reasons for Not Being Vaccinated

Among the study participants who have not been vaccinated against COVID-19, the differences between students' and parents' willingness to receive the vaccine in the 2 groups were not significant (Table S10 in Multimedia Appendix 1). After stratifying by parents' sex, the differences between the 2 groups were still not significant (Table S11 in Multimedia Appendix 1). For students who had not been vaccinated against COVID-19 after the intervention, the students and their parents were worried about side effects among many other reasons (Table S12 in Multimedia Appendix 1).

Discussion

Principal Results

Using a quasi-randomized controlled design, this study assessed the effectiveness of a video-assisted health education program sequenced by peer education on infectious disease health literacy among school students. The results suggest that the proposed multicomponent model of health education improved the knowledge of infectious diseases among students, and are consistent with those of previous randomized controlled trials in health education among primary school students [12,21,22]. Moreover, this study not only showcases an innovative approach to raise awareness of disease prevention by incorporating technology and behavioral elements, but also represents a preliminary effort to test the effectiveness of an infectious disease health education program using IRT-based scores.

Our results encapsulate important implications for the practice of health education and healthy behavior promotion. First, the inexpensive and convenient innovative health education approach proposed in this study represents a viable approach to improve student health literacy during pandemics and should be considered in future programs of healthy behavior promotion among school students. The fact that the program was effective among third-grade students does not restrict the potential of this approach since senior students are likely to capture the contents of the program better than third-grade students. Second, the results from the second session of this study partially indicate that web-based teaching may also be an effective tool to promote student engagement in health education, which has been highlighted in previous studies but not confirmed [7].

The possible long-term effects of the first session from our findings should not be ignored. The postintervention survey was carried out immediately after the second education session (including chicken pox, mumps, and COVID-19) and 4 months after the first education session (including influenza and norovirus diarrhea). Despite the time elapsed, the correct answer rates of questions related to infectious diseases that were of focus in the first session were still higher in the intervention group than in the control group. Therefore, third-grade primary school students may endure the impact of health education for at least 4 months. Given the low likelihood of frequently setting up health education sessions in schools, the slow waning of the program's effects is a desirable feature. However, the cross-over effect from the second session could not be ruled out. For example, the learning of COVID-19 may strengthen the students' previous understanding of influenza and increase the effect of intervention in influenza. In addition, the second session may sensitize the students in the intervention group. They may review the knowledge of the first session to prepare for the postintervention quiz, which may also enhance the effect of the first session.

XSL•FO
RenderX

It is noteworthy that there was some difference in response rates between the interventional and control groups. The difference in response rates might be attributable to an absence of treatment blinding. In fact, the intervention in this study could not be blinded due to its physical nature, in which case, the intervention group students might be motivated by the education sessions to meet the expectation of the educators to respond to the surveys.

In addition, there was no significant difference in the correct answer rate for questions related to flu transmission routes before or after the intervention, but the pre-post changes in the correct answer rates was different between the 2 groups, and the intervention group performed better than the control group. Owing to countrywide vaccination campaigns, the COVID-19 vaccination rates between the intervention and control groups were not significantly different. The results of the 2 questionnaire surveys showed that the vaccination rates of the 2 groups increased, which was related to the local epidemic and the country's policy encouragement for vaccination.

Limitations

Several limitations of the study should be noted when interpreting the results. First, we did not collect data on the incidence of related infectious diseases before and after the intervention. A previous study reported that in areas with a high incidence of infectious disease, the health education package had no overall effect in preventing infections. However, the intervention was effective in preventing infections in areas where the baseline prevalence was relatively low [21]. Further studies are needed to explore the impact of our composite intervention on preventing infections. Second, this study was limited in its ability to evaluate component-specific versus composite effects of the educational video, the didactic lessons, the cooperative learning exercises, and peer engagement. The 2-arm trial design could not parse out the influence of each element. Future work should incorporate multiple comparison arms to better isolate the impacts of intervention components. Third, we regret that we did not measure changes in attitudes and behaviors after the intervention, as the health education package is hypothesized to influence these aspects. This is a gap that exists in our study, which future research could explore. Fourth, we used a self-rating questionnaire to collect data. Although self-reporting is a common and accepted method, we could not completely rule out the possibility of measurement error. However, the reliability and validity of self-reporting among children aged >8 years have been shown to be good in health-related questionnaires [23,24]. Fifth, the contamination in this study may underestimate the effect of our intervention. We adopted a clustered quasi-randomized controlled trial design to mitigate within-class person-to-person contamination, although interclass contamination caused by students and teachers could not be eliminated. However, the contamination, if any, happened more likely to the first session rather than the second session since students were physically isolated during the latter. Sixth, as we did not receive the questionnaire from the students lost to follow-up, the primary analysis was not intent-to-treat. Seventh, the second session of health education originally scheduled to enter the campus was changed to web-based classes owing to the serious local epidemic. Therefore, the students were required to fill in the web-based questionnaire at home, which affected the independence of the participants in answering questions; hence, the correct answer rates of the 2 groups were generally higher than those at baseline. Besides, the recovery of the questionnaire was decreased probably due to the lack of the teachers' supervision outside the schools. However, the missing rates were balanced between the 2 groups, thereby reducing the chances of influencing our conclusions. Moreover, the effect of the health education provided herein may be underestimated because this missing group of students and parents might have lower health literacy, in which case, the intervention would have incremental value.

Comparison With Prior Work

Despite these limitations, the primary strengths of our study are that it is the first quasi-randomized controlled trial to evaluate the effect of a video-assisted health education program sequenced by peer education on the health literacy of COVID-19 and other infectious diseases among school children, and it is also the first to report IRT scores for questions related to the infectious diseases. Additionally, while our study is quasi-randomized, the allocation process likely achieved reasonable randomization, effectively balancing confounding factors across study arms as evidenced by the systematic allocation of students to intervention or control groups based on their odd or even class numbers, as outlined in Table 1. Importantly, the allocation of students to odd or even classes was not based on systematically different characteristics, as the Ministry of Education of the People's Republic of China does not permit students to be segregated into different classes based on specific attributes. Therefore, the grouping of students based on class number parity can be considered to approximate the effects of randomization. Moreover, the sample size in this study allowed minimal chances of underpowered analyses. Previous studies might have engaged nonrandomized designs such that mixed results were reported [12,21,22,25-30]. Although most studies demonstrated that the health intervention is effective in improving health knowledge and health literacy, a quasi-randomized controlled trial in China found that the intervention's effect was not significant among primary school students [25]. Moreover, a number of studies adopted self-control, or observational designs, based on which solid conclusions are difficult to derive [3-5,26-30].

Conclusions

Our study confirmed that the combination of video-assisted and peer education in a health education program had significant effects on school children. In addition, the effect of the first health education session may endure after 4 months. As such, the proposed program was effective in improving health literacy related to infectious diseases among school children and should be considered for en masse health promotion campaigns during pandemics.

XSL•FO
RenderX

Acknowledgments

This study was carried out by the Longgang District Bureau of Health as an administrative task. This sponsor was involved in developing the intervention and study design, and was kept informed during data collection and data analysis. The corresponding authors had full access to all of the data in the study and had the final responsibility for the decision to submit the paper for publication. We also appreciate Jing Chen and Wanjing Zhang from the Longgang District Health Inspection Institute for their contribution and support to the early development of the intervention package.

Data Availability

The data presented in this study are not publicly available because of ethical requirements but are accessible upon reasonable request to the corresponding author.

Authors' Contributions

All authors conceptualized and designed the study and provided administrative, technical, and material support; they also supervised the study. XZ and YJW acquired the data. XZ analyzed and interpreted the data and drafted the manuscript. YJ critically revised the manuscript for important intellectual content. XZ carried out the statistical analyses. NH is a co-correspondent (email: 207baby@163.com).

Conflicts of Interest

None declared.

Multimedia Appendix 1
Supplementary figures and tables.
[DOCX File , 922 KB - humanfactors_v11i1e43943_app1.docx]

Multimedia Appendix 2
CONSORT eHEALTH checklist (V 1.6.1).
[PDF File (Adobe PDF File), 1240 KB - humanfactors_v11i1e43943_app2.pdf]

References

1. Dong Y, Wang L, Burgner DP, Miller JE, Song Y, Ren X, et al. Infectious diseases in children and adolescents in China: analysis of national surveillance data from 2008 to 2017. BMJ 2020 Apr 02;369:m1043 [FREE Full text] [doi: 10.1136/bmj.m1043] [Medline: 32241761]
2. Wu Y, Liao Y, Zhou L, Luo Q, Wang Y, Li B. Monitoring of student absence due to illness in the first semester of 2018 in Shenzhen. Chin J Sch Health 2019;40(10):1584-1586. [doi: 10.16835/j.cnki.1000-9817.2019.10.044]
3. Qiu H, Yu Y, Liu B, Shen Y, Shao H. Study on effect of health education of hepatitis B among the primary, middle and high school students in Zhangjiagang of Jiangsu Province. J Med Pest Control 2019;35(08):751-747. [doi: 10.7629/yxdwfz201908008]
4. Liu J, Liu W. Evaluation of the effect of school-based sexuality education in reducing AIDS discrimination in primary school. Chin J Sch Health 2019;35(12):1121-1110. [doi: 10.16168/j.cnki.issn.1002-9982.2019.12.012]
5. Daoyong C, Shuo W, Tong L, Fei W, Yue L, Jixing Y. Effect evaluation of the video education on infectious disease knowledge among primary school students in Hongkou District, Shanghai. Health Promot Educ 2019;14(06):529-531. [doi: 10.16117/j.cnki.31-1974/r.201906018]
6. Layzer C, Rosapep L, Barr S. Student voices: perspectives on peer-to-peer sexual health education. J Sch Health 2017 Jul;87(7):513-523. [doi: 10.1111/josh.12519] [Medline: 28580671]
7. Coyne E, Rands H, Frommolt V, Kain V, Plugge M, Mitchell M. Investigation of blended learning video resources to teach health students clinical skills: an integrative review. Nurse Educ Today 2018 Apr;63:101-107 [FREE Full text] [doi: 10.1016/j.nedt.2018.01.021] [Medline: 29425738]
8. Li P, Li X, Meng H, Huang L, Zhang L, Wang S, et al. Video-assisted health education promotes rehabilitation training of total knee arthroplasty patients and reduces stress and burnout in nurses compared to oral education. Biomed Res Int 2021;2021:5058899 [FREE Full text] [doi: 10.1155/2021/5058899] [Medline: 34722765]
9. Seymour JE, Almack K, Kennedy S, Froggatt K. Peer education for advance care planning: volunteers' perspectives on training and community engagement activities. Health Expect 2013 Mar;16(1):43-55 [FREE Full text] [doi: 10.1111/j.1369-7625.2011.00688.x] [Medline: 21615641]
10. Benè KL, Bergus G. When learners become teachers: a review of peer teaching in medical student education. Fam Med 2014;46(10):783-787 [FREE Full text] [Medline: 25646829]

XSL•FO
RenderX

11. Mbachu C, Dim C, Ezeoke U. Effects of peer health education on perception and practice of screening for cervical cancer among urban residential women in south-east Nigeria: a before and after study. BMC Womens Health 2017 Jun 09;17(1):41 [FREE Full text] [doi: 10.1186/s12905-017-0399-6] [Medline: 28599635]
12. Tang Y, Diao H, Jin F, Pu Y, Wang H. The effect of peer education based on adolescent health education on the resilience of children and adolescents: A cluster randomized controlled trial. PLoS One 2022;17(2):e0263012 [FREE Full text] [doi: 10.1371/journal.pone.0263012] [Medline: 35108312]
13. Ajzen I. The theory of planned behavior. Organ Behav Hum Decis Process 1991 Dec;50(2):179-211. [doi: 10.1016/0749-5978(91)90020-T]
14. Frank L, Basch E, Selby JV, Patient-Centered Outcomes Research Institute. The PCORI perspective on patient-centered outcomes research. JAMA 2014 Oct 15;312(15):1513-1514. [doi: 10.1001/jama.2014.11100] [Medline: 25167382]
15. Yao Q, Li Y, Song L. Construction safety knowledge sharing on YouTube from 2007 to 2021: Two-step flow theory and semantic analysis. Safety Science 2022 Sep;153:105796. [doi: 10.1016/j.ssci.2022.105796]
16. Wright J. International Encyclopedia of the Social & Behavioral Sciences (second edition). Amsterdam: Elsevier; 2015.
17. Liu SH, Juster R, Dams-O'Connor K, Spicer J. Allostatic load scoring using item response theory. Compr Psychoneuroendocrinol 2021 Feb;5:100025 [FREE Full text] [doi: 10.1016/j.cpnec.2020.100025] [Medline: 35754455]
18. Fries JF, Bruce B, Cella D. The promise of PROMIS: using item response theory to improve assessment of patient-reported outcomes. Clin Exp Rheumatol 2005;23(5 Suppl 39):S53-S57. [Medline: 16273785]
19. Nguyen TH, Han H, Kim MT, Chan KS. An introduction to item response theory for patient-reported outcome measurement. Patient 2014 Jan 9;7(1):23-35 [FREE Full text] [doi: 10.1007/s40271-013-0041-0] [Medline: 24403095]
20. Rizopoulos D. ltm: an R Package for latent variable modeling and item response analysis. J Stat Softw 2007;17 [FREE Full text] [doi: 10.18637/jss.v017.i05]
21. Mationg MLS, Williams GM, Tallo VL, Olveda RM, Aung E, Alday P, et al. "The Magic Glasses Philippines": a cluster randomised controlled trial of a health education package for the prevention of intestinal worm infections in schoolchildren. Lancet Reg Health West Pac 2022 Jan;18:100312 [FREE Full text] [doi: 10.1016/j.lanwpc.2021.100312] [Medline: 35024651]
22. Hu P, Han L, Sharma M, Zeng H, Zhang Y, Li H, et al. Evaluation of cognitive and behavioral effects of peer education model-based intervention to sun safe in children. Iran J Public Health 2014 Mar;43(3):300-309 [FREE Full text] [Medline: 25988089]
23. Riley AW. Evidence that school-age children can self-report on their health. Ambul Pediatr 2004;4(4 Suppl):371-376. [doi: 10.1367/A03-178R.1] [Medline: 15264962]
24. Rebok G, Riley A, Forrest C, Starfield B, Green B, Robertson J, et al. Elementary school-aged children's reports of their health: a cognitive interviewing study. Qual Life Res 2001;10(1):59-70. [doi: 10.1023/a:1016693417166] [Medline: 11508476]
25. Xu L, Zhou C, Zhou J. Evaluation on effect of health promotion of school construction on health literacy of students in Changsha. Bull Dis Control Prev 2021;36(05):70-73. [doi: 10.13215/j.cnki.jbyfkztb.2107033]
26. Chen M, Wu C, Tsai C, Tsai M. Enterovirus board game for elementary school children: a pilot study. Public Health Nurs 2022 Mar;39(2):500-505. [doi: 10.1111/phn.12976] [Medline: 34537984]
27. Odone A, Bricchi L, Signorelli C. COVID-19 control school-based interventions: characteristics and impact of a national-level educational programme in Italy. Acta Biomed 2022 Jan 17;92(S6):e2021495 [FREE Full text] [doi: 10.23750/abm.v92iS6.12327] [Medline: 35037911]
28. Supriatun E, Nursasi AY, Fitriyani P. Enhancement of pulmonary tuberculosis prevention behavior with role play among elementary school students. Compr Child Adolesc Nurs 2017;40(sup1):78-87. [doi: 10.1080/24694193.2017.1386974] [Medline: 29166188]
29. Wang M, Han X, Fang H, Xu C, Lin X, Xia S, et al. Impact of health education on knowledge and behaviors toward infectious diseases among students in Gansu Province, China. Biomed Res Int 2018;2018:6397340 [FREE Full text] [doi: 10.1155/2018/6397340] [Medline: 29707573]
30. Taware S, Gawai P, Chatterjee A, Thakur H. Outcome of school-based intervention program in promoting personal hygiene in primary school children of Mumbai, India. Int Q Community Health Educ 2018 Oct;39(1):31-38. [doi: 10.1177/0272684X18809487] [Medline: 30376407]

Abbreviations

IRT: item response theory
TPM: three-parameter model

XSL•FO
RenderX

Edited by P Santana-Mancilla; submitted 31.10.22; peer-reviewed by W Ni, D Laborde, M Nasiri, Y Elsanousi; comments to author 28.07.23; revised version received 10.09.23; accepted 29.12.23; published 29.01.24.

Please cite as:
Zhang X, Wen YJ, Han N, Jiang Y
The Effect of a Video-Assisted Health Education Program Followed by Peer Education on the Health Literacy of COVID-19 and Other Infectious Diseases Among School Children: Quasi-Randomized Controlled Trial
JMIR Hum Factors 2024;11:e43943
URL: https://humanfactors.jmir.org/2024/1/e43943
doi:10.2196/43943
PMID:38285496

©Xiaojuan Zhang, Yingkun Justin Wen, Ning Han, Yawen Jiang. Originally published in JMIR Human Factors (https://humanfactors.jmir.org), 29.01.2024. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on https://humanfactors.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Barriers to and Facilitators of Key Stakeholders Influencing Successful Digital Implementation of Remote Monitoring Solutions: Mixed Methods Analysis

Fahad Mujtaba Iqbal[1], BSc, MBChB, PhD; Ravi Aggarwal[1], PhD; Meera Joshi[1], PhD; Dominic King[1], MBChB; Guy Martin[1], PhD; Sadia Khan[2], MBChB; Mike Wright[3], PhD; Hutan Ashrafian[1], PhD; Ara Darzi[1], FMedSci

[1]Division of Surgery, Imperial College London, London, United Kingdom

[2]West Middlesex University Hospital, London, United Kingdom

[3]Innovation Business Partner, Chelsea and Westminster NHS Trust, London, United Kingdom

Corresponding Author:
Fahad Mujtaba Iqbal, BSc, MBChB, PhD
Division of Surgery
Imperial College London
10th Floor QEQM Building
Praed Street
London, W2 1NY
United Kingdom
Phone: 44 020 3312 6666
Email: fahad.iqbal@doctors.org.uk

Abstract

Background: Implementation of remote monitoring solutions and digital alerting tools in health care has historically been challenging, despite the impetus provided by the COVID-19 pandemic. To date, a health systems–based approach to systematically describe barriers and facilitators across multiple domains has not been undertaken.

Objective: We aimed to undertake a comprehensive mixed methods analysis of barriers and facilitators for successful implementation of remote monitoring and digital alerting tools in complex health organizations.

Methods: A mixed methods approach using a modified Technology Acceptance Model questionnaire and semistructured interviews mapped to the validated fit among humans, organizations, and technology (HOT-fit) framework was undertaken. Likert frequency responses and deductive thematic analyses were performed.

Results: A total of 11 participants responded to the questionnaire and 18 participants to the interviews. Key barriers and facilitators could be mapped onto 6 dimensions, which incorporated aspects of digitization: system use (human), user satisfaction (human), environment (organization), structure (organization), information and service quality (technology), and system quality (technology).

Conclusions: The recommendations proposed can enhance the potential for future remote sensing solutions to be more successfully integrated in health care practice, resulting in more successful use of "virtual wards."

Trial Registration: ClinicalTrials.gov NCT05321004; https://www.clinicaltrials.gov/study/NCT05321004

(JMIR Hum Factors 2024;11:e49769) doi:10.2196/49769

KEYWORDS

implementation science; health plan implementation; mobile health; health care industry; stakeholder; COVID-19; remote monitoring; digital tools; digital health; pandemic; virtual wards; virtual ward; health care delivery; telemedicine; telehealth; wearables; wearable; technology; United Kingdom; UK; digital services

Introduction

Since the onset of the COVID-19 pandemic, adoption and implementation of novel health care pathways have accelerated globally. A key change has been transitioning beyond the traditional face-to-face model of health care delivery with the incorporation of novel remote monitoring solutions [1,2]. They offer a significant advantage in moderating viral exposure risk

XSL•FO
RenderX

to health care staff, reducing community spread, and delivering quality health care remotely for exposed or infected individuals [3,4].

The integration of telemedicine and remote monitoring into medical practice is expected to expand by appropriately permitting selected individuals to continue living at home rather than admitting them into secondary care; this very premise is the foundation of "virtual wards" [5]. With the recent improvements made to wearable technology, they can support health provider assessment and clinical decision-making through collected biometric data both in secondary care and in the community [6-10].

However, successful implementation of digital technologies across complex hospital systems is seldom a smooth process [11-13]. The absence of standardized procedures for implementation and evaluation alongside the deficiency of published implementation strategies adds to these difficulties. One study in the National Health Service (NHS) that implemented wearable sensors and alerting systems in secondary care reported no improvements in clinical outcomes among patients [14,15]. The aim was to use wearable sensors to provide continuous remote monitoring to patients admitted to acute (nonintensive) wards and alert health care staff upon recognition of deterioration. Interestingly, although the digital solution was able to pick up clinical deterioration in vital signs and alert health care staff, responding to the alert was met with significant delay. This was in spite of health care staff in the NHS reporting favorable perceptions of digital solutions with potential improvements to patient safety and reduced staff burden [16]. Therefore, there is a need to further explore implementation issues.

Patients have reported high levels of acceptance, comfort, and safety and deemed such digital tools favorable [17-19]. The main concerns, from a patient perspective, surround potential overreliance on numbers with diminishing contact from clinical staff [17,20,21]. Health care staff perceptions, however, have been more mixed, with concerns regarding changing and increasing workloads, uncertainty surrounding the clinical meaningfulness of captured data, and alert fatigue [19-21]. Although mixed methods exploration of these 2 key stakeholder groups has been well documented, understanding how to integrate remote monitoring digital tools in the NHS requires further examination of cultural and management issues in the health care organization, an area where evidence is missing.

In the United Kingdom, large health informatics programs and widespread digital transformations are delivered by NHS Digital, a nondepartmental public body [22,23]. To support digitization, NHS England has formed a framework consisting of 3 ambitions: digital readiness, maturity, and data-enabled services [24]. In line with this, NHS England has supported the development and use of virtual wards, further indicating the "digital push" [5]. For policy makers, understanding the barriers and facilitators as perceived by key organizational members is crucial for the effective provision and smooth deployment of digitally enabled care. A proposed framework evaluates these aspects, incorporating the concept of fit among humans, organizations, and technology (HOT-fit) [25]. This framework offers a structured basis to examine factors that focus on alignment and compatibility across these 3 domains, thereby enhancing the effectiveness of digital health care initiatives.

Therefore, the aim of this study was to evaluate key stakeholder perspectives on an organizational level of implementing remote monitoring solutions in the NHS, identifying factors that could affect successful execution and adoption using the HOT-fit framework. In doing so, we propose a road map for implementing wearable solutions in secondary care.

Methods

Study Design

A mixed methods approach was implemented that consisted of semistructured interviews and questionnaires [26]. This was developed in accordance with recommendations from the Standards for Reporting Qualitative Research (SRQR) guidelines where appropriate [27]. The semistructured interviews were conducted with high-level stakeholders from industry and academia, as well as with health care providers who played an instrumental role in and had prior experience of implementing digital solutions. Additionally, a validated questionnaire was used to ascertain the perceived technological acceptance of new remote monitoring systems.

To ensure appropriate recruitment among all key stakeholder groups, a key informant strategy was followed for purposive recruitment [28,29]. Individuals were identified through their notable work with implementation of remote monitoring solutions in health care, including authors of impactful research in the literature, major digital technology companies, technicians involved with digital tool infrastructure development, and experts recommended by peers. This represented a variety of groups, including academics, clinicians, allied health care professionals, and employees of Google Health, who had experience with implementing digital solutions with the NHS.

Ethical Considerations

All recruited participants provided written informed consent. Ethical approval for this study was obtained by Imperial College London's Science Engineering Technology Research Ethics Committee (20IC6331), and it was conducted in accordance with the Good Clinical Practice guidelines and the Declaration of Helsinki. Storage and handling of personal data complied with the General Data Protection Regulation. Interviews were recorded, anonymized, and transcribed.

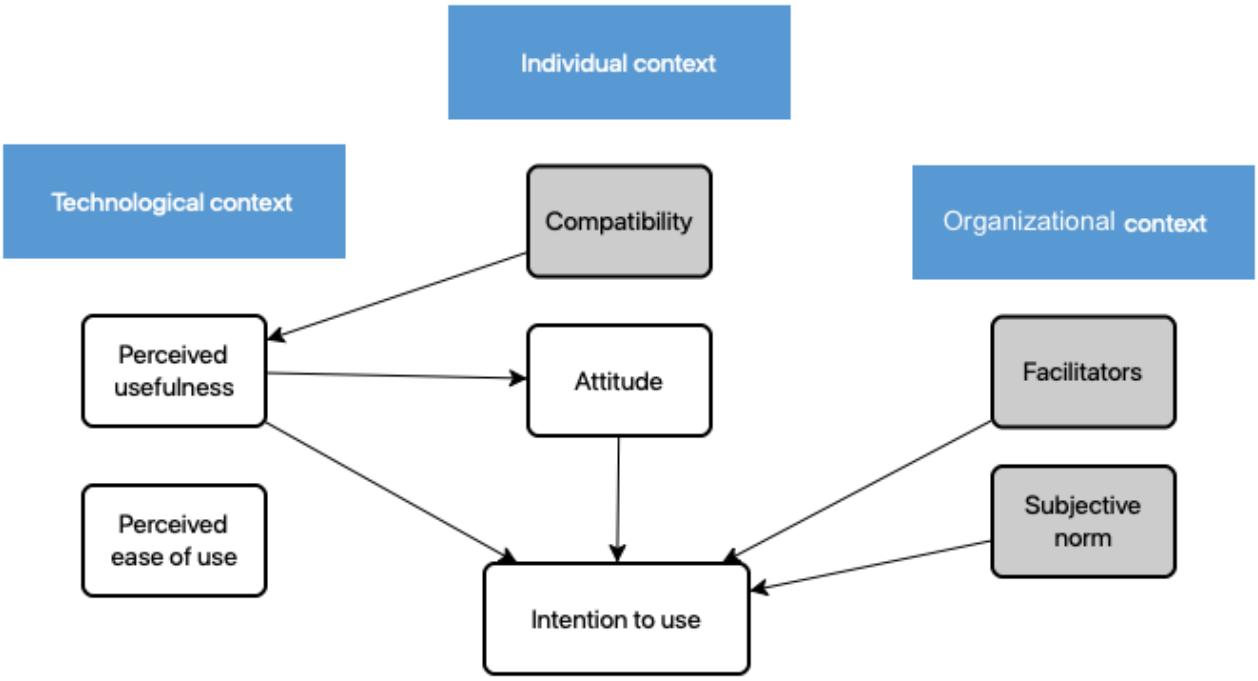
Questionnaires

An adapted version of the Technology Acceptance Model (TAM) questionnaire was used; this validated questionnaire has shown acceptably high Cronbach α values [30]. This ensures the reliability of our findings, contributing to the robustness of the study's methodology and its implications in understanding technology acceptance dynamics. The proposed theoretical framework (information technology acceptance) is shown in Figure 1. It has been adapted from Chau and Hu [31], comprising individual context, technological context, and organizational context. Further adaptations from Gagnon et al [30], with the inclusion of theories of interpersonal behavior

XSL•FO
RenderX

and reasoned action building on the TAM, proposed by Davis [32], have been included [30-34]. As such, individual context consists of compatibility (factors that affect acceptance of a new technology) and attitude (perception of the individual to adopting a technology); technological context consists of perceived usefulness and perceived ease of use of technologies. Lastly, organizational context consists of facilitators and subjective norms; the latter can be described as social (an individual's perception of a behavior) or descriptive (behavior of others).

Figure 1. Theoretical framework for the modified Technology Acceptance Model questionnaire [30].



Semistructured Interviews

All participants were invited to take part in semistructured interviews conducted by the lead researchers. A structured topic guide was created following a literature review that drew heavily from a model proposed by Simblett et al [35] and by the HOT-fit framework [25].

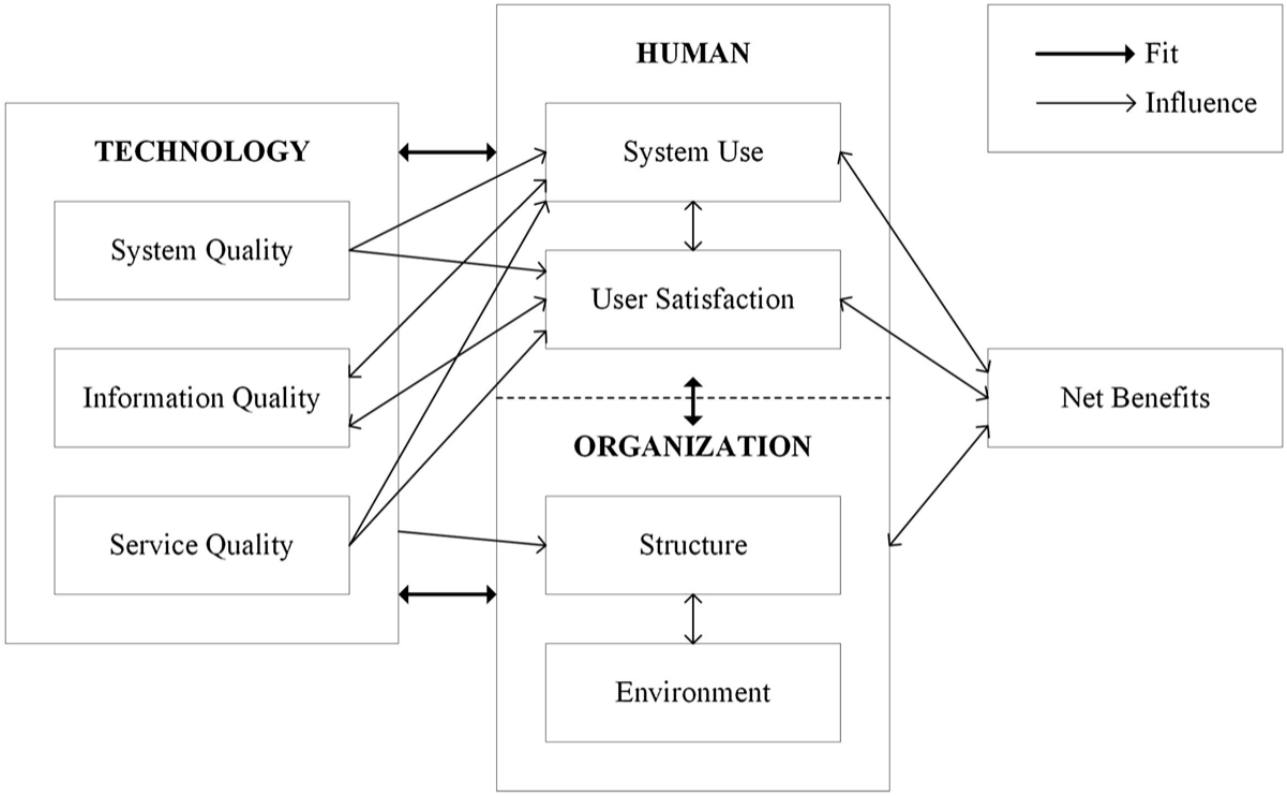
Data collection was an iterative process; emerging recurring concepts were incorporated into the interview guide for further exploration with remaining participants. Interviews were recorded, anonymized, and transcribed verbatim before being entered into NVivo (version 12; QSR International) for analysis.

HOT-Fit Framework

This validated framework identifies dimensions that can be mapped onto and used as reference models for evaluating the performance, effectiveness, and impact of health systems [25,36]. A fit between human, organizational, and technological factors is required to ensure successful implementation and has been highlighted in Figure 2.

XSL•FO
RenderX

Figure 2. The fit among humans, organizations, and technology (HOT-fit) framework, adapted from Yusof et al [25].



Data Analysis

Frequency distributions were generated for the 7-point Likert scale responses to the modified TAM questionnaire using R studio (R Foundation for Statistical Computing) with the *Likert* package (Bryer and Speerschneider).

Transcribed interviews were analyzed using a broadly deductive approach [37], with the topic guide adapted as previously described [35]. This formed the basis for the initial predefined coding framework and was undertaken by 2 independent researchers to determine barriers and facilitators [37]. An iterative process of coding and data indexing occurred, ensuring key aspects were not missed from the predefined coding framework. Subsequent emerging themes were summarized and mapped to the evaluation measures corresponding to each dimension of the HOT-fit framework [25]. The results were discussed until consensus was reached.

Results

Overviews of the included participants and the reported evaluation measures are shown in Tables 1 and 2, respectively.

Table 1. Demographics of included participants.

Group	Role 1	Role 2	Role 3	Role 4	Role 5	Role 6	Role 7
Health care trusts	Director of strategy, research and innovation	Chief clinical information officer and Caldicott Guardian	Digital quality improvement lead	Project manager	Chief information officer	Systems, integration interoperability architect	Lead nurse for remote monitoring
Academics	Clinical lecturer	Clinical lecturer	Chief scientific advisor	—[a]	—	—	—
Google Health	Clinical lead	Clinical specialist	Product manager	Implementation specialist	Implementation manager	Program manager	—
Other	Programme director: innovation of health	Managing director: digital health	—	—	—	—	—

[a]—Not applicable.

XSL•FO
RenderX

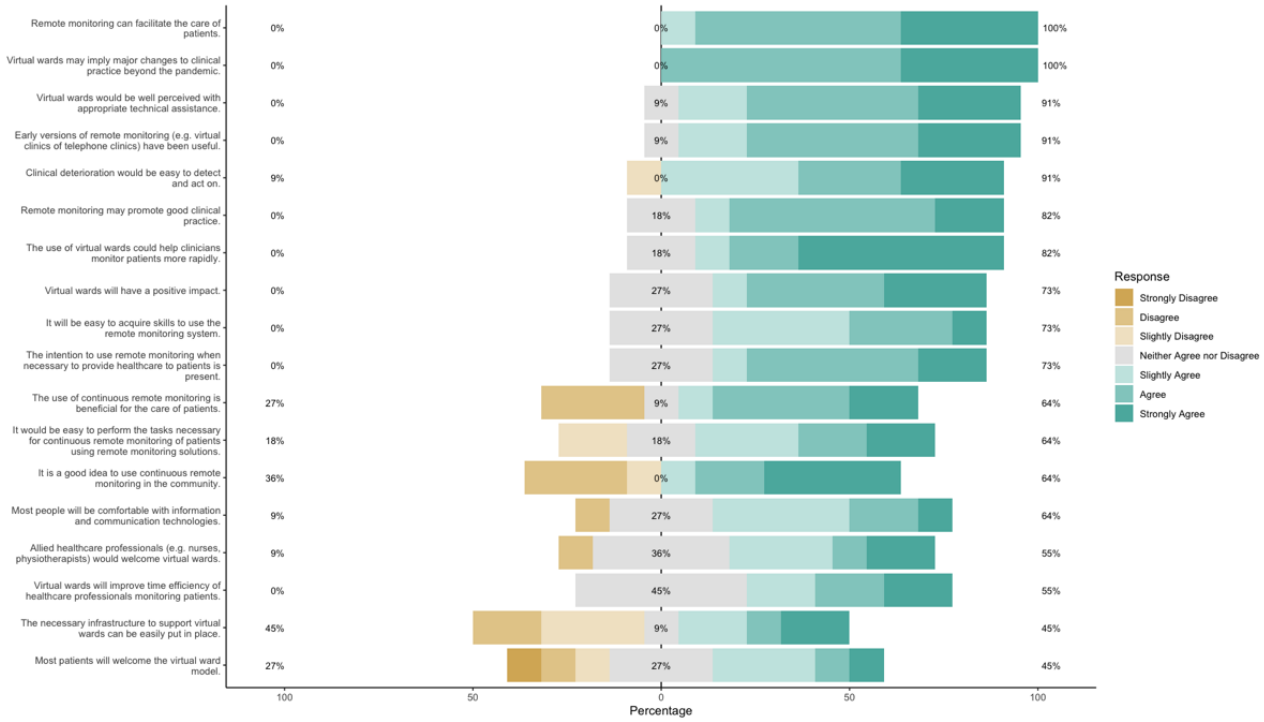
Table 2. Overview of reported evaluation measures.

Dimension and evaluation measures	Factors
System use	
Expectation and beliefs	• Improved efficiency (facilitator) • Appropriate selection of end users suitable for digital tool (facilitator)
Training, knowledge, and expertise	• Lack of troubleshooting support (barrier) • Engagement with new starters (facilitator)
Motivation	• Large data burden (barrier) • Post–COVID-19 fatigue of staff (barrier) • Finding local champions (facilitator)
User satisfaction (no evaluation measures)	• Developing relationships for feedback (facilitator) • Previous negative experiences with no feedback on benefit (barrier)
Environment (no evaluation measures)	• Overburdened National Health Service system (barrier)
Structure (clinical process)	• Clear strategic framework and partnership (facilitator)
Information and service quality (no evaluation measures)	• Poor interoperability (barrier) • Poor user interface and user engagement (barrier)
System quality	• Failure to provide added value (barrier)

TAM Questionnaire

A total of 11 participants (response rate 11/22, 50%) responded to the questionnaire; the responses are represented as a Likert plot (Figure 3). Overall, the technology surrounding remote monitoring and virtual wards was perceived well by the questioned stakeholders, who considered that it facilitated the care of patients and that these pathways, initially introduced during the pandemic, were likely to change long-term provision of health care. However, some concerns were noted regarding whether the existing infrastructure could support the technology's use and whether it would improve efficiency. Of note, there was uncertainty regarding whether most patients would welcome virtual wards or remote monitoring.

Figure 3. Likert plot displaying responses to the modified Technology Acceptance Model questionnaire. The percentages on the left and right sides of the plot represent the totals for negative and positive responses, while the percentages in the center represent neutral responses.



XSL•FO
RenderX

Semistructured Interviews

A total of 22 participants were approached, of whom 18 (response rate: 82%) participated in the semistructured interviews (Table 1). An overview of the factors, by dimension, that respondents felt were responsible for contributing to implementation is summarized in Table 2.

System Use

Expectations and Beliefs

The prospect of introducing novel remote monitoring technologies was felt to facilitate implementation through improved efficiency, particularly since the implementation of electronic health records has improved data availability and clarity:

> *...with the implementation [and] introduction of electronic health records where the data that's available is so granular. And in addition to new technologies that are coming. I think that you can do a lot more, remotely or virtually, and it does make things a lot more efficient...* [Participant 15]

Moreover, respondents also commented that for successful implementation, a selective process should be in place for patients who would benefit the most from novel technologies, rather than using the technology in cases that would not be meaningful:

> *From a patient perspective, we don't want to one size fits all approach. We need to be clear about how we personalize this and how it's relevant and meaningful.* [Participant 17]

Training, Knowledge, and Expertise

Problems with troubleshooting and available training were reported to reduce successful implementation due to a lack of support:

> *We've had problems when trying to use the remote monitoring, it came up with an error and then I have to try and sort that out, you know? It's just things like that that make extra work.* [Participant 16]

> *I know that the nurses have struggled a huge amount with remote monitoring, and I expected that...because there's a lot of upskilling.* [Participant 18]

However, engaging early with health care workers and obtaining their involvement was shown to improve implementation of remote monitoring solutions:

> *[We received] better engagement by tying the implementation with the new starters in the role and the changeover of junior doctors, because it was a new product to offer to new junior doctors.* [Participant 3]

Motivation

It was felt that motivation to engage with technologies would be impacted through the excessive availability of data acting as a deterrent:

> *We need to be mindful about the data burdens, not just for patients but for staff because this kind of remote technology follows you around. You basically could work 24/7 365 of the year.* [Participant 17]

In addition, following the pandemic, many health care workers were fatigued and unmotivated to engage in change, acting as a barrier to successful remote technology implementation:

> *Post-COVID the workforce has been decimated, been exhausted and is fatigued. It's not the only problem though, because you know as well as I do that the NHS has run this model of where it's good will. We've never had infrastructure that we needed to do stuff and we still get a huge amount done. So it's not the only driver at the moment. It's more noticeable because of where people's heads are at and obviously where their physical levels and mental levels of exhaustion are...* [Participant 17]

However, respondents also noted that finding a few motivated individuals to champion change at a local level can help implementation:

> *I asked them to self-nominate three of them who were interested in helping [implement]. So they led and supported the [technology]...* [Participant 18]

User Satisfaction

Respondents reported that previous experience with digital tools tied into user satisfaction. Feedback to end users demonstrating meaningful impact was deemed important for engagement and successful implementation:

> *Where staff or patients, for example, have been involved in projects before that they haven't had any feedback from, haven't seen any meaningful outcome from...they're like, well, why would I want to get engaged with this? That's a lot of energy and effort from me and I won't see any benefit.* [Participant 17]

> *...develop relationships, so between, if you like, supplier and developer and clinical staff so you've got these rapid cycles of feedback and learning.* [Participant 1]

Environment

Respondents reported that previous hindrance of effective implementation was because of an overburdened system unable to give the appropriate attention to integrating a digital solution in the NHS:

> *NHS is overburdened and so that level of diligence...wasn't there until it had to be, until things became mission critical...that comes down to a bandwidth problem...* [Participant 10]

Similarly, underresourcing was noted to be a barrier, particularly during the early stages, where issues would arise:

> *More resource[s] to get [things] kick started [are usually needed]...because we had to go through all the teething problems ourselves which created extra work for us.* [Participant 16]

> *We've got very limited resources, that they're very thinly spread across all of the IT projects that require integration and interoperability...just the sheer volume*

XSL•FO
RenderX

> *of work that the Trust has heaped on us over the last three or four years is the bigger constraining factor.* [Participant 4]

Lastly, organizational culture supporting digitization was a commonly reported theme, with some institutions more readily accepting of innovation than others:

> *Organisational culture can be both the barrier and facilitator. We know that there are some organisations that are much more ready and able to adopt innovation. I think from an organisational perspective, competing priorities are a huge issue...If your IT is majorly engaged in doing something else, for example an EHR implementation, its ability to support remote monitoring and other technologies is really poor.* [Participant 17]

Structure

Respondents also commented on the need for a clear process and said that developing a strategic partnership and framework would facilitate implementation and should be planned before rollout:

> *Strategic framework is crucial on things.... What does a strategic partnership look like? What is the direction that we want to jointly head in? What do we want to achieve together, and what are the different components to get there...* [Participant 1]

> *Making [the product vision and roadmap] clear as early on and getting that input right at the beginning of any kind of feature development. So that there is expectation alignment on what is being developed whether the minimal viable product meets the use cases that it needs to, and that there's a partnership in prioritizing these features and when they're delivered. As opposed to just showing a feature set a few weeks before it gets deployed.... I think that initial understanding of the vision...and getting that clinical engagement as early on helps to set the path going forward.* [Participant 12]

Information and Service Quality

Respondents noted the need for digital tools to be interoperable and usable, as poorly designed digital tools would be a barrier, hindering an overly strained NHS system:

> *The challenges are IT and interoperability…you don't want 20 bits of data…from 20 apps that don't work, so that's the usability and the accessibility and the staffing of these models because traditionally they basically get added onto someone's day job. But that day person's already overwhelmed.* [Participant 17]

System Quality

Respondents highlighted that for a digital tool to be successfully implemented, it needed to provide added value, with perceived usefulness and ease of use being crucial.

> *[What] was the added value in [this digital app]? All it did was render some of the information that we already had in a limited manner, back in the mobile device.* [Participant 8]

> *Usability, the accessibility, and the staffing of these models [are really poor] because traditionally they basically get added onto someone's day job.... The data element [is also] really poor, so you get a lot of enthusiasts doing a lot of projects. But if you then say where's your evidence that makes any difference to anything meaningful that matters to patients and staff, they can't produce that. I think the digital health tech industry has been really slow at that.* [Participant 17]

Furthermore, it was believed that the best way to implement a digital tool (eg, remote monitoring solutions) was through rapid quality improvement cycles following the plan-do-study-act (PDSA) technique, focusing on targeting user experience issues:

> *...believe the technology suffered from very poor clinical and user engagement. So I know [technological companies] will tell us they've had loads of user engagement, but actually most patients wouldn't say that, they'd say well, why is it like this? No, why is nobody been engaged in the design for this?* [Participant 17]

> *...trying to give clinical input into feasibility, usability, implementation in terms of the design of how we were going to implement stuff, so...[a] genuine PDSA type approach to implementation, and I was quite involved in some of the thinking about spread and how do you get this utilized across different parts of the Trust...* [Participant 1]

Discussion

Principal Findings

This study explored barriers and facilitators for implementing digital tools, in particular remote monitoring solutions, in the NHS, alongside the acceptance of such technology using the modified TAM questionnaire. Using the HOT-fit framework, human, organization, and technological factors were categorized, allowing for a multiple-angled approach to a multifaceted problem. Therefore, key barriers and facilitators could be mapped onto 6 dimensions, which incorporated aspects of digitization: system use (human), user satisfaction (human), environment (organization), structure (organization), information and service quality (technology), and system quality (technology).

With regards to system use, the importance of improving workflow efficiency, having appropriate troubleshooting support available for staff, finding local champions to help integration within the clinical workforce, and positively engaging with health care staff were highlighted as facilitators. To support this, young staff have been deemed the most likely to engage with and benefit from a new workflow [36,38-40]. This, in part, may be explained by more adept digital literacy skills and technical proficiencies associated with junior members [41]. In the literature, concise and tailored education surrounding implementation has been promoted as an important facilitator [42].

XSL•FO
RenderX

Key barriers relating to system use and environment included poor training and the burden of data, particularly with continuous remote monitoring of vital signs. These data may not always be clinically meaningful or because of poor resourcing may not be acknowledged appropriately, generating additional work for existing staff, who are already overburdened [14,43]. Previously, this unincentivized workflow change led to poor response times to alerts generated through alerting systems in an acute surgical ward [14]. In this study, 36% (4/11) of respondents to the modified TAM questionnaire were unsure whether allied health care professionals would welcome virtual wards (Figure 3). One study highlighted that these workers, in particular nurses and clinicians, were the most important gatekeepers for remote monitoring solutions [44]. Therefore, engaging these groups, fostering positive relationships, and delivering regular feedback would enhance user satisfaction, allow user interface and engagement issues to be proactively tackled, and subsequently enable successful implementation.

Concerning system quality, perceived usefulness and ease of use were deemed as important facilitators for successful implementation. In the literature, intuitive and user-friendly systems have been confirmed to have easier acceptance [36,39]. The modified TAM questionnaire similarly confirmed this in our cohort, particularly through questions concerning acquiring new skills and impact, emphasizing that remote monitoring technology could be readily accepted.

Limitations

This study included key stakeholders belonging to a broad selection of groups (academics, industry, and health care) in order to create a broad understanding of factors that influence implementation of remote monitoring solutions in the NHS. Given that previous studies have focused on end user testing, this study sought to provide a top-down view to give a better understanding of considerations that could influence widespread implementation [16,18]. However, in doing so, our interpretations have some limitations. First, the broad, heterogeneous sample of key stakeholders included may identify issues that are generalizable, but the nonprobabilistic sampling may have resulted in a selection bias. Moreover, the included sample size was limited. Despite this, the use of semistructured interviews yielded pertinent considerations for pragmatic implementation in hospital settings. In addition, differences between various hospitals and departments, which may have different attitudes toward digital technologies, were not explored in this study. A final limitation relates to the HOT-fit framework; although it is considered useful, the mapping of factors is a subjective undertaking and mapping to one specific measure was, at times, difficult.

Further Research and Recommendations

Although our cohort showed that there was overall acceptance of remote monitoring technology (Figure 3), there remains a deficiency with respect to successful implementation. This was noted most recently in one study where the median time to acknowledge an alert from health care staff was 111 (range 1-2146) minutes, despite early recognition of deterioration from remote sensing [14]. Therefore, further research should incorporate human factors and behavior evaluation when implementing remote monitoring solutions with the NHS; moreover, implementation frameworks such as HOT-fit should be used to ensure multiple angles have been carefully considered.

To facilitate the effective integration of remote monitoring solutions in clinical workflows, a comprehensive strategic framework is paramount. This framework should prioritize the early involvement of end users, fostering relationships that enable rapid feedback on implementation strategies, user interfaces, and user experience issues. Such engagement allows for iterative enhancements through PDSA cycles, promoting continuous improvement [45].

Industries aiming to develop remote monitoring technologies must collaborate closely with key stakeholders, ensuring the creation of products that provide significant value and feature user-friendly interfaces. This approach emphasizes the importance of a bottom-up strategy in technology implementation, valuing the autonomy and insights of end users, who play a crucial role in the successful adoption of these solutions. Crucial to this process is the establishment of robust infrastructural support prior to the deployment of remote monitoring systems. Adequate resourcing and the involvement of technical support staff are essential to facilitate seamless integration with existing information technology frameworks, thereby enhancing the prioritization and effectiveness of digital health initiatives. By adhering to these guidelines, health care organizations can enhance the integration of remote monitoring into clinical practice, leading to improved operational efficiency, patient care, and overall health care service delivery.

Conclusion

Implementation of remote monitoring solutions in the NHS remains a complex challenge. The results of this study have highlighted key stakeholder perceptions that could influence successful integration. Through the proposed recommendations, there is potential for future remote sensing solutions to be more successfully integrated into our health care practices, resulting in novel pathways expanding beyond virtual wards.

Acknowledgments

Infrastructure support for this research was provided by the National Institute for Health and Care Research (NIHR) Imperial Biomedical Research Centre and the NIHR Imperial Patient Safety Translational Research Centre.

Data Availability

The data sets used and/or analyzed during the study are available from the corresponding author on reasonable request.

XSL•FO
RenderX

Authors' Contributions

FMI drafted the manuscript. Significant amendments were made by RA, MJ, MW, SK, HA, and AD. All authors approved the final manuscript.

Conflicts of Interest

AD is chair of the Health Security initiative and HA is chief scientific officer at Flagship Pioneering UK Ltd. Flagship Pioneering had no role in the development, conduct, or analysis of the study.

References

1. Watson AR, Wah R, Thamman R. The value of remote monitoring for the COVID-19 pandemic. Telemed J E Health 2020 Sep;26(9):1110-1112. [doi: 10.1089/tmj.2020.0134] [Medline: 32384251]
2. The impact of the COVID-19 pandemic on outpatient visits: changing patterns of care in the newest COVID-19 hot spots. The Commonwealth Fund. URL: https://www.commonwealthfund.org/publications/2020/aug/impact-covid-19-pandemic-outpatient-visits-changing-patterns-care-newest [accessed 2020-09-19]
3. Xu S, Li Y. Beware of the second wave of COVID-19. Lancet 2020 Apr 25;395(10233):1321-1322 [FREE Full text] [doi: 10.1016/S0140-6736(20)30845-X] [Medline: 32277876]
4. WHO coronavirus disease (COVID-19) dashboard. World Health Organization. URL: https://covid19.who.int/ [accessed 2020-06-25]
5. Supporting information: virtual ward including hospital at home. NHS England. URL: https://www.england.nhs.uk/publication/supporting-information-virtual-ward-including-hospital-at-home/ [accessed 2022-06-07]
6. Martirosyan M, Caliskan K, Theuns D, Szili-Torok T. Remote monitoring of heart failure: benefits for therapeutic decision making. Expert Rev Cardiovasc Ther 2017 Jul;15(7):503-515. [doi: 10.1080/14779072.2017.1348229] [Medline: 28656875]
7. Annis T, Pleasants S, Hultman G, Lindemann E, Thompson JA, Billecke S, et al. Rapid implementation of a COVID-19 remote patient monitoring program. J Am Med Inform Assoc 2020 Aug 01;27(8):1326-1330 [FREE Full text] [doi: 10.1093/jamia/ocaa097] [Medline: 32392280]
8. Ford D, Harvey JB, McElligott J, King K, Simpson KN, Valenta S, et al. Leveraging health system telehealth and informatics infrastructure to create a continuum of services for COVID-19 screening, testing, and treatment. J Am Med Inform Assoc 2020 Dec 09;27(12):1871-1877 [FREE Full text] [doi: 10.1093/jamia/ocaa157] [Medline: 32602884]
9. O'Carroll O, MacCann R, O'Reilly A, Dunican EM, Feeney ER, Ryan S, et al. Remote monitoring of oxygen saturation in individuals with COVID-19 pneumonia. Eur Respir J 2020 Aug;56(2):2001492 [FREE Full text] [doi: 10.1183/13993003.01492-2020] [Medline: 32616588]
10. Medina M, Babiuch C, Card M, Gavrilescu R, Zafirau W, Boose E, et al. Home monitoring for COVID-19. Cleve Clin J Med 2020 Jun 11:1-4 [FREE Full text] [doi: 10.3949/ccjm.87a.ccc028] [Medline: 32409432]
11. Wachter R. Making IT work: harnessing the power of health information technology to improve care in England. Gov.UK. URL: https://assets.publishing.service.gov.uk/media/5a8091afe5274a2e87dba8f2/Wachter_Review_Accessible.pdf [accessed 2023-12-02]
12. Greenhalgh T, Swinglehurst D, Stones R. Rethinking resistance to 'big IT': a sociological study of why and when healthcare staff do not use nationally mandated information and communication technologies. Health Soc Care Deliv Res 2014 Nov;2(39):1-112. [doi: 10.3310/hsdr02390] [Medline: 27466649]
13. Kruse C, Betancourt J, Ortiz S, Valdes Luna SM, Bamrah IK, Segovia N. Barriers to the use of mobile health in improving health outcomes in developing countries: systematic review. J Med Internet Res 2019 Oct 09;21(10):e13263 [FREE Full text] [doi: 10.2196/13263] [Medline: 31593543]
14. Iqbal F, Joshi M, Fox R, Koutsoukou T, Sharma A, Wright M, et al. Outcomes of vital sign monitoring of an acute surgical cohort with wearable sensors and digital alerting systems: a pragmatically designed cohort study and propensity-matched analysis. Front Bioeng Biotechnol 2022;10:895973 [FREE Full text] [doi: 10.3389/fbioe.2022.895973] [Medline: 35832414]
15. Iqbal FM, Joshi M, Khan S, Ashrafian H, Darzi A. Implementation of wearable sensors and digital alerting systems in secondary care: protocol for a real-world prospective study evaluating clinical outcomes. JMIR Res Protoc 2021 May 04;10(5):e26240 [FREE Full text] [doi: 10.2196/26240] [Medline: 33944790]
16. Joshi M, Archer S, Morbi A, Ashrafian H, Arora S, Khan S, et al. Perceptions on the use of wearable sensors and continuous monitoring in surgical patients: interview study among surgical staff. JMIR Form Res 2022 Feb 11;6(2):e27866 [FREE Full text] [doi: 10.2196/27866] [Medline: 35147503]
17. Downey C, Brown J, Jayne D, Randell R. Patient attitudes towards remote continuous vital signs monitoring on general surgery wards: an interview study. Int J Med Inform 2018 Jun;114:52-56. [doi: 10.1016/j.ijmedinf.2018.03.014] [Medline: 29673603]
18. Joshi M, Archer S, Morbi A, Arora S, Kwasnicki R, Ashrafian H, et al. Short-term wearable sensors for in-hospital medical and surgical patients: mixed methods analysis of patient perspectives. JMIR Perioper Med 2021 Apr 22;4(1):e18836 [FREE Full text] [doi: 10.2196/18836] [Medline: 33885367]

XSL•FO
RenderX

19. Leenen JPL, Dijkman EM, van Dijk JD, van Westreenen HL, Kalkman C, Schoonhoven L, et al. Feasibility of continuous monitoring of vital signs in surgical patients on a general ward: an observational cohort study. BMJ Open 2021 Feb 17;11(2):e042735 [FREE Full text] [doi: 10.1136/bmjopen-2020-042735] [Medline: 33597138]
20. Weenk M, van Goor H, Frietman B, Engelen LJ, van Laarhoven CJ, Smit J, et al. Continuous monitoring of vital signs using wearable devices on the general ward: pilot study. JMIR Mhealth Uhealth 2017 Jul 05;5(7):e91 [FREE Full text] [doi: 10.2196/mhealth.7208] [Medline: 28679490]
21. Areia C, King E, Ede J, Young L, Tarassenko L, Watkinson P, et al. Experiences of current vital signs monitoring practices and views of wearable monitoring: A qualitative study in patients and nurses. J Adv Nurs 2022 Mar;78(3):810-822 [FREE Full text] [doi: 10.1111/jan.15055] [Medline: 34655093]
22. About NHS Digital. National Health Service. URL: https://digital.nhs.uk/about-nhs-digital [accessed 2020-09-19]
23. NHS Digital annual report and accounts 2019 to 2020. National Health Service. URL: https://tinyurl.com/2pcn7y6b [accessed 2020-09-19]
24. Digital framework for allied health professionals. NHS England. 2019. URL: https://www.england.nhs.uk/wp-content/uploads/2019/04/a-digital-framework-for-allied-health-professionals.pdf [accessed 2023-12-02]
25. Yusof M, Kuljis J, Papazafeiropoulou A, Stergioulas L. An evaluation framework for health information systems: human, organization and technology-fit factors (HOT-fit). Int J Med Inform 2008 Jun;77(6):386-398. [doi: 10.1016/j.ijmedinf.2007.08.011] [Medline: 17964851]
26. Iqbal FM, Joshi M, Khan S, Wright M, Ashrafian H, Darzi A. Key stakeholder barriers and facilitators to implementing remote monitoring technologies: protocol for a mixed methods analysis. JMIR Res Protoc 2022 Jul 21;11(7):e38437 [FREE Full text] [doi: 10.2196/38437] [Medline: 35862185]
27. O'Brien BC, Harris IB, Beckman TJ, Reed DA, Cook DA. Standards for reporting qualitative research: a synthesis of recommendations. Acad Med 2014 Sep;89(9):1245-1251 [FREE Full text] [doi: 10.1097/ACM.0000000000000388] [Medline: 24979285]
28. Marshall MN. The key informant technique. Fam Pract 1996 Feb;13(1):92-97. [doi: 10.1093/fampra/13.1.92] [Medline: 8671109]
29. Campbell S, Greenwood M, Prior S, Shearer T, Walkem K, Young S, et al. Purposive sampling: complex or simple? Research case examples. J Res Nurs 2020 Dec;25(8):652-661 [FREE Full text] [doi: 10.1177/1744987120927206] [Medline: 34394687]
30. Gagnon MP, Orruño E, Asua J, Abdeljelil AB, Emparanza J. Using a modified technology acceptance model to evaluate healthcare professionals' adoption of a new telemonitoring system. Telemed J E Health 2012;18(1):54-59 [FREE Full text] [doi: 10.1089/tmj.2011.0066] [Medline: 22082108]
31. Chau P, Hu P. Examining a model of information technology acceptance by individual professionals: an exploratory study. J Manag Inf Syst 2014 Dec 23;18(4):191-229. [doi: 10.1080/07421222.2002.11045699]
32. Davis FD. Perceived usefulness, perceived ease of use, and user acceptance of information technology. MIS Q 1989 Sep;13(3):319. [doi: 10.2307/249008]
33. Hill RJ, Fishbein M, Ajzen I. Belief, attitude, intention and behavior: an introduction to theory and research. Contemp Sociol 1977 Mar;6(2):244. [doi: 10.2307/2065853]
34. Triandis HC. Nebraska Symposium on Motivation, 1979: Beliefs, Attitudes and Values. Lincoln, NB: University of Nebraska Press; 1980:195-259.
35. Simblett S, Greer B, Matcham F, Curtis H, Polhemus A, Ferrão J, et al. Barriers to and facilitators of engagement with remote measurement technology for managing health: systematic review and content analysis of findings. J Med Internet Res 2018 Jul 12;20(7):e10480 [FREE Full text] [doi: 10.2196/10480] [Medline: 30001997]
36. Kilsdonk E, Peute LW, Jaspers MWM. Factors influencing implementation success of guideline-based clinical decision support systems: A systematic review and gaps analysis. Int J Med Inform 2017 Feb;98:56-64. [doi: 10.1016/j.ijmedinf.2016.12.001] [Medline: 28034413]
37. Braun V, Clarke V. Using thematic analysis in psychology. Qual Res Psychol 2006 Jan;3(2):77-101. [doi: 10.1191/1478088706qp063oa]
38. Lugtenberg M, Pasveer D, van der Weijden T, Westert GP, Kool RB. Exposure to and experiences with a computerized decision support intervention in primary care: results from a process evaluation. BMC Fam Pract 2015 Oct 16;16(1):141 [FREE Full text] [doi: 10.1186/s12875-015-0364-0] [Medline: 26474603]
39. Varonen H, Kortteisto T, Kaila M, EBMeDS Study Group. What may help or hinder the implementation of computerized decision support systems (CDSSs): a focus group study with physicians. Fam Pract 2008 Jun;25(3):162-167. [doi: 10.1093/fampra/cmn020] [Medline: 18504253]
40. Zheng K, Padman R, Johnson MP, Diamond HS. Understanding technology adoption in clinical care: clinician adoption behavior of a point-of-care reminder system. Int J Med Inform 2005 Aug;74(7-8):535-543. [doi: 10.1016/j.ijmedinf.2005.03.007] [Medline: 16043083]
41. Mobasheri MH, King D, Johnston M, Gautama S, Purkayastha S, Darzi A. The ownership and clinical use of smartphones by doctors and nurses in the UK: a multicentre survey study. BMJ Innov 2015 Oct 07;1(4):174-181. [doi: 10.1136/BMJINNOV-2015-000062]

XSL•FO
RenderX

42. Trivedi MH, Daly EJ, Kern JK, Grannemann BD, Sunderajan P, Claassen CA. Barriers to implementation of a computerized decision support system for depression: an observational report on lessons learned in "real world" clinical settings. BMC Med Inform Decis Mak 2009 Jan 21;9:6 [FREE Full text] [doi: 10.1186/1472-6947-9-6] [Medline: 19159458]
43. Downey C, Randell R, Brown J, Jayne D. Continuous versus intermittent vital signs monitoring using a wearable, wireless patch in patients admitted to surgical wards: pilot cluster randomized controlled trial. J Med Internet Res 2018 Dec 11;20(12):e10802 [FREE Full text] [doi: 10.2196/10802] [Medline: 30538086]
44. Whitten PS, Mackert MS. Addressing telehealth's foremost barrier: provider as initial gatekeeper. Int J Technol Assess Health Care 2005;21(4):517-521. [doi: 10.1017/S0266462305050725] [Medline: 16262977]
45. Leis JA, Shojania KG. A primer on PDSA: executing plan-do-study-act cycles in practice, not just in name. BMJ Qual Saf 2017 Jul;26(7):572-577. [doi: 10.1136/bmjqs-2016-006245] [Medline: 27986900]

Abbreviations

HOT-fit: fit among humans, organizations, and technology
NHS: National Health Service
PDSA: plan-do-study-act
SRQR: Standards for Reporting Qualitative Research
TAM: Technology Acceptance Model

Edited by A Kushniruk; submitted 08.06.23; peer-reviewed by A Acharya, C Sakal; comments to author 19.01.24; revised version received 26.01.24; accepted 07.04.24; published 06.05.24.

Please cite as:
Iqbal FM, Aggarwal R, Joshi M, King D, Martin G, Khan S, Wright M, Ashrafian H, Darzi A
Barriers to and Facilitators of Key Stakeholders Influencing Successful Digital Implementation of Remote Monitoring Solutions: Mixed Methods Analysis
JMIR Hum Factors 2024;11:e49769
URL: https://humanfactors.jmir.org/2024/1/e49769
doi:10.2196/49769
PMID:37338929

©Fahad Mujtaba Iqbal, Ravi Aggarwal, Meera Joshi, Dominic King, Guy Martin, Sadia Khan, Mike Wright, Hutan Ashrafian, Ara Darzi. Originally published in JMIR Human Factors (https://humanfactors.jmir.org), 06.05.2024. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on https://humanfactors.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Publisher:
JMIR Publications
130 Queens Quay East.
Toronto, ON, M5A 3Y5
Phone: (+1) 416-583-2040
Email: support@jmir.org

XSL•FO
RenderX